PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification:
C12N 15/57, C07K 14/47, C07K 16/18, C07K 19/00, C12N 1/21, C12N 5/10, C12N 9/64, C12N 15/12, C12N 15/62, C12N 15/85, C12Q 1/37, G01N 33/68

A2

(11) International Publication Number: WO 00/17369

(43) International Publication Date: 30 March 2000 (30.03.2000)

(21) International Application Number: PCT/US99/20881

(22) International Filing Date: 23 September 1999 (23.09.1999)

(30) Priority Data:
60/101,594 24 September 1998 (24.09.1998) US

(60) Parent Application or Grant
PHARMACIA & UPJOHN COMPANY [/]; (). GURNEY, Mark, E. [/]; (). BIENKOWSKI, Michael, Jerome [/]; (). HEINRIKSON, Robert, Leroy [/]; (). PARODI, Luis, A. [/]; (). YAN, Riqiang [/]; (). GURNEY, Mark, E. [/]; (). BIENKOWSKI, Michael, Jerome [/]; (). HEINRIKSON, Robert, Leroy [/]; (). PARODI, Luis, A. [/]; (). YAN, Riqiang [/]; (). WOOTTON, Thomas, A. ; ().

Published

(54) Title: ALZHEIMER'S DISEASE SECRETASE

(54) Titre: SECRETASE DE LA MALADIE D'ALZHEIMER

(57) Abstract

The present invention provides the enzyme and enzymatic procedures for cleaving the 'beta' secretase cleavage site of the APP protein and associated nucleic acids, peptides, vectors, cells and cell isolates and assays.

(57) Abrégé

La présente invention porte sur l'enzyme et les procédures enzymatiques de clivage du site de clivage de la 'beta' secrétase de la protéine APP et des acides nucléiques, des peptides, des vecteurs, des cellules et des isolats cellulaires associés, et sur des dosages.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/57, 15/62, 15/85, 5/10, 9/64, C07K 19/00, 14/47, C12N 15/12, C07K 16/18, C12Q 1/37, G01N 33/68, C12N 1/21

A2

(11) International Publication Number: WO 00/17369

(43) International Publication Date: 30 March 2000 (30.03.00)

(21) International Application Number: PCT/US99/20881

(22) International Filing Date: 23 September 1999 (23.09.99)

(30) Priority Data:
60/101,594 24 September 1998 (24.09.98) US

(71) Applicant *(for all designated States except US)*: PHARMACIA & UPJOHN COMPANY [US/US]; 301 Henrietta Street, Kalamazoo, MI 49001 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: GURNEY, Mark, E. [US/US]; 910 Rosewood Avenue, S.E., Grand Rapids, MI 49506 (US). BIENKOWSKI, Michael, Jerome [US/US]; 3431 Hollow Wood, Portage, MI 49024 (US). HEINRIKSON, Robert, Leroy [US/US]; 81 South Lake Doster Drive, Plainwell, MI 49080 (US). PARODI, Luis, A. [US/SE]; Grevgafan 24, S-115 43 Stockholm (SE). YAN, Riqiang [US/US]; 5026 Queen Victoria Street, Kalamazoo, MI 49009 (US).

(74) Agent: WOOTTON, Thomas, A.; Pharmacia & Upjohn Company, Intellectual Property Legal Services, 301 Henrietta Street, Kalamazoo, MI 49001 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: ALZHEIMER'S DISEASE SECRETASE

(57) Abstract

The present invention provides the enzyme and enzymatic procedures for cleaving the β secretase cleavage site of the APP protein and associated nucleic acids, peptides, vectors, cells and cell isolates and assays.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

Description

Alzheimer's Disease Secretase

FIELD OF THE INVENTION

The present invention related to the field of Alzheimer's Disease, APP, amyloid beta peptide, and human aspartyl proteases as well as a method for the identification of agents that modulate the activity of these polypeptides.

BACKGROUND OF THE INVENTION

Alzheimer's disease (AD) causes progressive dementia with consequent formation of amyloid plaques, neurofibrillary tangles, gliosis and neuronal loss. The disease occurs in both genetic and sporadic forms whose clinical course and pathological features are quite similar. Three genes have been discovered to date which when mutated cause an autosomal dominant form of Alzheimer's disease. These encode the amyloid protein precursor (APP) and two related proteins, presenilin-1 (PS1) and presenilin-2 (PS2), which as their names suggest are both structurally and functionally related. Mutations in any of the three enhance proteolytic processing of APP via an intracellular pathway that produces amyloid beta peptide or the Aβ peptide (or sometimes here as Abeta), a 40-42 amino acid long peptide that is the primary component of amyloid plaque in AD. Dysregulation of intracellular pathways for proteolytic processing may be central to the pathophysiology of AD. In the case of plaque formation, mutations in APP, PS1 or PS2 consistently alter the proteolytic processing of APP so as to enhance formation of Aβ 1-42, a form of the Aβ peptide which seems to be particularly amyloidogenic, and thus very important in AD. Different forms of APP range in size from 695-770 amino acids, localize to the cell surface, and have a single C-terminal transmembrane domain. The Abeta peptide is derived from a region of APP adjacent to and containing a portion of the transmembrane domain. Normally, processing of APP at the α-secretase site cleaves the midregion of the Aβ sequence adjacent to the membrane and releases the soluble, extracellular domain of APP from the cell surface. This α-secretase APP processing, creates soluble APP- α, and it is normal and not thought to contribute to AD.

Pathological processing of APP at the β- and γ-secretase sites produces a very different result than processing at the α site. Sequential processing at the β- and γ-secretase sites releases the Aβ peptide, a peptide possibly very important in AD pathogenesis. Processing at the β- and γ-secretase sites can occur in both the endoplasmic reticulum (in neurons) and in the endosomal/lysosomal pathway after reinternalization of cell surface

APP (in all cells). Despite intense efforts, for 10 years or more, to identify the enzymes responsible for processing APP at the β and γ sites, to produce the Aβ peptide, those proteases remained unknown until this disclosure. Here, for the first time, we report the identification and characterization of the β secretase enzyme. We disclose some known and some novel human aspartic proteases that can act as β-secretase proteases and, for the first time, we explain the role these proteases have in AD. We describe regions in the proteases critical for their unique function and for the first time characterize their substrate. This is the first description of expressed isolated purified active protein of this type, assays that use the protein, in addition to the identification and creation of useful cell lines and inhibitors.

SUMMARY OF THE INVENTION

Here we disclose a number of variants of the asp2 gene and peptide.

Any isolated or purified nucleic acid polynucleotide that codes for a protease capable of cleaving the beta (β) secretase cleavage site of APP that contains two or more sets of special nucleic acids, where the special nucleic acids are separated by nucleic acids that code for about 100 to 300 amino acid positions, where the amino acids in those positions may be any amino acids, where the first set of special nucleic acids consists of the nucleic acids that code for the peptide DTG, where the first nucleic acid of the first special set of nucleic acids is, the first special nucleic acid, and where the second set of nucleic acids code for either the peptide DSG or DTG, where the last nucleic acid of the second set of nucleic acids is the last special nucleic acid, with the proviso that the nucleic acids disclosed in SEQ ID NO. 1 and SEQ. ID NO. 5 are not included. The nucleic acid polynucleotide of claim 1 where the two sets of nucleic acids are separated by nucleic acids that code for about 125 to 222 amino acid positions, which may be any amino acids.The nucleic acid polynucleotide of claim 2 that code for about 150 to 172 amino acid positions, which may be any amino acids. The nucleic acid polynucleotide of claim that code for about 172 amino acid positions, which may be any amino acids. The nucleic acid polynucleotide of claim 4 where the nucleotides are described in SEQ. ID. NO. 3 The nucleic acid polynucleotide of claim 2 where the two sets of nucleic acids are separated by nucleic acids that code for about 150 to 196 amino acid positions. The nucleic acid polynucleotide of claim 6 where the two sets of nucleotides are separated by nucleic acids that code for about 196 amino acids (positions). The nucleic acid polynucleotide of claim 7 where the two sets of nucleic acids are separated by the same nucleic acid sequences that separate the same set of special nucleic acids in SEQ. ID. NO. 5. The nucleic acid

polynucleotide of claim 4 where the two sets of nucleic acids are separated by nucleic acids that code for about 150 to 190, amino acid (positions). The nucleic acid polynucleotide of claim 9 where the two sets of nucleotides are separated by nucleic acids that code for about 190 amino acids (positions). The nucleic acid polynucleotide of claim 10 where the two sets of nucleotides are separated by the same nucleic acid sequences that separate the same set of special nucleotides in SEQ. ID. NO. 1. Claims 1-11 where the first nucleic acid of the first special set of amino acids, that is, the first special nucleic acid, is operably linked to any codon where the nuclic acids of that codon codes for any peptide comprising from 1 to 10,000 amino acid (positions). The nucleic acid polynucleotide of claims 1-12 where the first special nucleic acid is operably linked to nucleic acid polymers that code for any peptide selected from the group consisting of: any any reporter proteins or proteins which facilitate purification.The nucleic acid polynucleotide of claims 1-13 where the first special nucleic acid is operably linked to nucleic acid polymers that code for any peptide selected from the group consisting of: immunoglobin-heavy chain, maltose binding protein, glutathion S transfection, Green Fluorescent protein, and ubiquitin. Claims 1-14 where the last nucleic acid of the second set of special amino acids, that is, the last special nucleic acid, is operably linked to nucleic acid polymers that code for any peptide comprising any amino acids from 1 to 10,000 amino acids. Claims 1-15 where the last special nucleic acid is operably linked to any codon linked to nucleic acid polymers that code for any peptide selected from the group consisting of: any reporter proteins or proteins which facilitate purification. The nucleic acid polynucleotide of claims 1-16 where the first special nucleic acid is operably linked to nucleic acid polymers that code for any peptide selected from the group consisting of: immunoglobin-heavy chain, maltose binding protein, glutathion S transfection, Green Fluorescent protein, and ubiquitin.

Any isolated or purified nucleic acid polynucleotide that codes for a protease capable of cleaving the beta secretase cleavage site of APP that contains two or more sets of special nucleic acids, where the special nucleic acids are separated by nucleic acids that code for about 100 to 300 amino acid positions, where the amino acids in those positions may be any amino acids, where the first set of special nucleic acids consists of the nucleic acids that code for DTG, where the first nucleic acid of the first special set of nucleic acids is, the first special nucleic acid, and where the second set of nucleic acids code for either DSG or DTG, where the last nucleic acid of the second set of special nucleic acids is the last special nucleic acid, where the first special nucleic acid is operably linked to nucleic

acids that code for any number of amino acids from zero to 81 amino acids and where each of those codons may code for any amino acid. The nucleic acid polynucleotide of claim 18 , where the first special nucleic acid is operably linked to nucleic acids that code for any number of from 64 to 77 amino acids where each codon may code for any amino acid. The nucleic acid polynucleotide of claim 19 , where the first special nucleic acid is operably linked to nucleic acids that code for 71 amino acids. The nucleic acid polynucleotide of claim 20, where the first special nucleic acid is operably linked to 71 amino acids and where the first of those 71 amino acids is the amino acid T. The nucleic acid polynucleotide of claim 21, where the polynucleotide comprises a sequence that is at least 95% identical to SEQ. ID. (Example 11). The nucleic acid polynucleotide of claim 22, where the complete polynucleotide comprises SEQ. ID. (Example 11). The nucleic acid polynucleotide of claim 18 , where the first special nucleic acid is operably linked to nucleic acids that code for any number of from 40 to 54 amino acids where each codon may code for any amino acid. The nucleic acid polynucleotide of claim 24, where the first special nucleic acid is operably linked to nucleic acids that code for 47 amino acids. The nucleic acid polynucleotide of claim 20, where the first special nucleic acid is operably linked to 47 codons where the first those 47 amino acids is the amino acid E. The nucleic acid polynucleotide of claim 21, where the polynucleotide comprises a sequence that is at least 95% identical to SEQ. ID. (Example 10). The nucleic acid polynucleotide of claim 22, where the complete polynucleotide comprises SEQ. ID. (Example 10).

Any isolated or purified nucleic acid polynucleotide that codes for a protease capable of cleaving the beta (β) secretase cleavage site of APP that contains two or more sets of special nucleic acids, where the special nucleic acids are separated by nucleic acids that code for about 100 to 300 amino acid positions, where the amino acids in those positions may be any amino acids, where the first set of special nucleic acids consists of the nucleic acids that code for the peptide DTG, where the first nucleic acid of the first special set of amino acids is, the first special nucleic acid, and where the second set of special nucleic acids code for either the peptide DSG or DTG, where the last nucleic acid of the second set of special nucleic acids, the last special nucleic acid, is operably linked to nucleic acids that code for any number of codons from 50 to 170 codons. The nucleic acid polynucleotide of claim 29 where the last special nucleic acid is operably linked to nucleic acids comprising from 100 to 170 codons. The nucleic acid polynucleotide of claim 30 where the last special nucleic acid is operably linked to nucleic acids comprising from 142

to 163 codons. The nucleic acid polynucleotide of claim 31 where the last special nucleic acid is operably linked to nucleic acids comprising about 142 codons. The nucleic acid polynucleotide of claim 32 where the polynucleotide comprises a sequence that is at least 95% identical to SEQ. ID. (Example 9 or 10). The nucleic acid polynucleotide of claim 33, where the complete polynucleotide comprises SEQ. ID. (Example 9 or 10). The nucleic acid polynucleotide of claim 31 where the last special nucleic acid is operably linked to nucleic acids comprising about 163 codons. The nucleic acid polynucleotide of claim 35 where the polynucleotide comprises a sequence that is at least 95% identical to SEQ. ID. (Example 9 or 10). The nucleic acid polynucleotide of claim 36, where the complete polynucleotide comprises SEQ. ID. (Example 9 or 10). The nucleic acid polynucleotide of claim 31 where the last special nucleic acid is operably linked to nucleic acids comprising about 170 codons. Claims 1-38 where the second set of special nucleid acids code for the peptide DSG, and optionally the first set of nucleic acid polynucleotide is operably linked to a peptide purification tag. Claims 1-39 where the nucleic acid polynucleotide is operably linked to a peptide purification tag which is six histidine. Claims 1-40 where the first set of special nucleic acids are on one polynucleotide and the second set of special nucleic acids are on a second polynucleotide, where both first and second polynucleotides have at lease 50 codons. Claims 1-40 where the first set of special nucleic acids are on one polynucleotide and the second set of special nucleic acids are on a second polynucleotide, where both first and second polynucleotides have at lease 50 codons where both said polynucleotides are in the same solution. A vector which contains a polynucleotide described in claims 1-42. A cell or cell line which contans a polynucleotide described in claims 1-42.

Any isolated or purified peptide or protein comprising an amino acid polymer that is a protease capable of cleaving the beta (β) secretase cleavage site of APP that contains two or more sets of special amino acids, where the special amino acids are separated by about 100 to 300 amino acid positions, where each amino acid position can be any amino acid, where the first set of special amino acids consists of the peptide DTG, where the first amino acid of the first special set of amino acids is, the first special amino acid, where the second set of amino acids is selected from the peptide comprising either DSG or DTG, where the last amino acid of the second set of special amino acids is the last special amino acid, with the proviso that the proteases disclosed in SEQ ID NO. 2 and SEQ. ID NO. 6 are not included. The amino acid polypeptide of claim 45 where the two sets of amino acids are

separated by about 125 to 222 amino acid positions where in each position it may be any amino acid. The amino acid polypeptide of claim 46 where the two sets of amino acids are separated by about 150 to 172 amino acids. The amino acid polypeptide of claim 47 where the two sets of amino acids are separated by about 172 amino acids. The amino acid polypeptide of claim 48 where the protease is described in SEQ. ID. NO. 4 The amino acid polypeptide of claim 46 where the two sets of amino acids are separated by about 150 to 196 amino acids. The amino acid polypeptide of claim 50 where the two sets of amino acids are separated by about 196 amino acids. The amino acid polypeptide of claim 51 where the two sets of amino acids are separated by the same amino acid sequences that separate the same set of special amino acids in SEQ. ID. NO. 6. The amino acid polypeptide of claim 46 where the two sets of amino acids are separated by about 150 to 190, amino acids. The amino acid polypeptide of claim 53 where the two sets of nucleotides are separated by about 190 amino acids. The amino acid polypeptide of claim 54 where the two sets of nucleotides are separated by the same amino acid sequences that separate the same set of special amino acids in SEQ. ID. NO. 2. Claims 45-55 where the first amino acid of the first special set of amino acids, that is, the first special amino acid, is operably linked to any peptide comprising from 1 to 10,000 amino acids. The amino acid polypeptide of claims 45-56 where the first special amino acid is operably linked to any peptide selected from the group consisting of: any any reporter proteins or proteins which facilitate purification. The amino acid polypeptide of claims 45-57 where the first special amino acid is operably linked to any peptide selected from the group consisting of: immunoglobin-heavy chain, maltose binding protein, glutathion S transfection, Green Fluorescent protein, and ubiquitin. Claims 45-58, where the last amino acid of the second set of special amino acids, that is, the last special amino acid, is operably linked to any peptide comprising any amino acids from 1 to 10,000 amino acids. Claims 45-59 where the last special amino acid is operably linked any peptide selected from the group consisting of any reporter proteins or proteins which facilitate purification. The amino acid polypeptide of claims 45-60 where the first special amino acid is operably linked to any peptide selected from the group consisting of: immunoglobin-heavy chain, maltose binding protein, glutathion S transfection, Green Fluorescent protein, and ubiquitin.

Any isolated or purified peptide or protein comprising an amino acid polypeptide that codes for a protease capable of cleaving the beta secretase cleavage site of APP that contains two or more sets of special amino acids, where the special amino acids are

separated by about 100 to 300 amino acid positions, where each amino acid in each position can be any amino acid, where the first set of special amino acids consists of the amino acids DTG, where the first amino acid of the first special set of amino acids is, the first special amino acid, D, and where the second set of amino acids is either DSG or DTG, where the last amino acid of the second set of special amino acids is the last special amino acid, G, where the first special amino acid is operably linked to amino acids that code for any number of amino acids from zero to 81 amino acid positions where in each position it may be any amino acid. The amino acid polypeptide of claim 62, where the first special amino acid is operably linked to a peptide from about 64 to 77 amino acids positions where each amino acid position may be any amino acid. The amino acid polypeptide of claim 63, where the first special amino acid is operably linked to a peptide of 71 amino acids. The amino acid polypeptide of claim 64, where the first special amino acid is operably linked to 71 amino acids and the first of those 71 amino acids is the amino acid T. The amino acid polypeptide of claim 65, where the polypeptide comprises a sequence that is at least 95% identical to SEQ. ID. (Example 11). The amino acid polypeptide of claim 66, where the complete polypeptide comprises SEQ. ID. (Example 11). The amino acid polypeptide of claim 62, where the first special amino acid is operably linked to any number of from 40 to 54 amino acids (positions) where each amino acid position may be any amino acid. The amino acid polypeptide of claim 68, where the first special amino acid is operably linked to amino acids that code for a peptide of 47 amino acids. The amino acid polypeptide of claim 69, where the first special amino acid is operably linked to a 47 amino acid peptide where the first those 47 amino acids is the amino acid E. The amino acid polypeptide of claim 70, where the polypeptide comprises a sequence that is at least 95% identical to SEQ. ID. (Example 10). The amino acid polypeptide where the polypeptide comprises Example 10).

Any isolated or purified amino acid polypeptide that is a protease capable of cleaving the beta (β) secretase cleavage site of APP that contains two or more sets of special amino acids, where the special amino acids are separated by about 100 to 300 amino acid positions, where each amino acid in each position can be any amino acid, where the first set of special amino acids consists of the amino acids that code for DTG, where the first amino acid of the first special set of amino acids is, the first special amino acid, D, and where the second set of amino acids are either DSG or DTG, where the last amino acid of the second set of special amino acids is the last special amino acid, G, which is operably linked to any number of amino acids from 50 to 170 amino acids, which may be any amino

acids. The amino acid polypeptide of claim 73 where the last special amino acid is operably linked to a peptide of about 100 to 170 amino acids. The amino acid polypeptide of claim 74 where the last special amino acid is operably linked to to a peptide of about 142 to 163 amino acids. The amino acid polypeptide of claim 75 where the last special amino acid is operably linked to to a peptide of about about 142 amino acids. The amino acid polypeptide of claim 76 where the polypeptide comprises a sequence that is at least 95% identical to SEQ. ID. (Example 9 or 10). The amino acid polypeptide of claim 75 where the last special amino acid is operably linked to a peptide of about 163 amino acids. The amino acid polypeptide of claim 79 where the polypeptide comprises a sequence that is at least 95% identical to SEQ. ID. (Example 9 or 10). The amino acid polypeptide of claim 79, where the complete polypeptide comprises SEQ. ID. (Example 9 or 10). The amino acid polypeptide of claim 74 where the last special amino acid is operably linked to to a peptide of about 170 amino acids. Claim 46-81 where the second set of special amino acids is comprised of the peptide with the amino acid sequence DSG. Claims 45-82 where the amino acid polypeptide is operably linked to a peptide purification tag. Claims 45-83 where the amino acid polypeptide is operably linked to a peptide purification tag which is six histidine. Claims 45-84 where the first set of special amino acids are on one polypeptide and the second set of special amino acids are on a second polypeptide, where both first and second polypeptide have at lease 50 amino acids, which may be any amino acids. Claims 45-84 where the first set of special amino acids are on one polypeptide and the second set of special amino acids are on a second polypeptide, where both first and second polypeptides have at lease 50 amino acids where both said polypeptides are in the same vessel. A vector which contains a polypeptide described in claims 45-86. A cell or cell line which contans a polynucleotide described in claims 45-87. The process of making any of the polynucleotides, vectors, or cells of claims 1-44. The process of making any of the polypeptides, vectors or cells of claims 45-88. Any of the polynucleotides, polypeptides, vectors, cells or cell lines described in claims 1-88 made from the processes described in claims 89 and 90.

Any isolated or purified peptide or protein comprising an amino acid polypeptide that codes for a protease capable of cleaving the beta secretase cleavage site of APP that contains two or more sets of special amino acids, where the special amino acids are separated by about 100 to 300 amino acid positions, where each amino acid in each position can be any amino acid, where the first set of special amino acids consists of the amino acids

DTG, where the first amino acid of the first special set of amino acids is, the first special amino acid, D, and where the second set of amino acids is either DSG or DTG, where the last amino acid of the second set of special amino acids is the last special amino acid, G, where the first special amino acid is operably linked to amino acids that code for any number of amino acids from zero to 81 amino acid positions where in each position it may be any amino acid.

The amino acid polypeptide of claim 62, where the first special amino acid is operably linked to a peptide from about 30 to 77 amino acids positions where each amino acid position may be any amino acid. The amino acid polypeptide of claim 63, where the first special amino acid is operably linked to a peptide of 35, 47, 71, or 77 amino acids.

The amino acid polypeptide of claim 63, where the first special amino acid is operably linked to the same corresponding peptides from SEQ. ID. NO. 3 that are 35, 47, 71, or 77 peptides in length, beginning counting with the amino acids on the first special sequence, DTG, towards the N-terminal of SEQ. ID. NO. 3.

The amino acid polypeptide of claim 65, where the polypeptide comprises a sequence that is at least 95% identical to the same corresponding amino acids in SEQ. ID. NO. 4, that is, identical to that portion of the sequences in SEQ.ID. NO. 4, including all the sequences from both the first and or the second special nucleic acids, toward the N-terminal, through and including 71, 47, 35 amino acids before the first special amino acids. (Examples 10 and 11).

The amino acid polypeptide of claim 65, where the complete polypeptide comprises the peptide of 71 amino acids, where the first of the amino acid is T and the second is Q. The nucleic acid polynucleotide of claim 21, where the polynucleotide comprises a sequence that is at least 95% identical to the same corresponding amino acids in SEQ. ID. NO. 3, that is, identical to the sequences in SEQ. ID. NO. 3 including the sequences from both the first and or the second special nucleic acids, toward the N-Terminal, through and including 71 amino acids, see Example 10, beginning from the DTG site and including the nucleotides from that code for 71 amino acids).

The nucleic acid polynucleotide of claim 22, where the complete polynucleotide comprises identical to the same corresponding amino acids in SEQ. ID. NO. 3, that is, identical to the sequences in SEQ. ID. NO. 3 including the sequences from both the first and or the second special nucleic acids, toward the N-Terminal, through and including 71

amino acids, see Example 10, beginning from the DTG site and including the nucleotides from that code for 71 amino acids).

The nucleic acid polynucleotide of claim 18 , where the first special nucleic acid is operably linked to nucleic acids that code for any number of from about 30 to 54 amino acids where each codon may code for any amino acid.

The nucleic acid polynucleotide of claim 20, where the first special nucleic acid is operably linked to 47 codons where the first those 35 or 47 amino acids is the amino acid E or G.

The nucleic acid polynucleotide of claim 21, where the polynucleotide comprises a sequence that is at least 95% identical to the same corresponding amino acids in SEQ. ID. NO. 3, that is, identical to that portion of the sequences in SEQ. ID. NO. 3 including the sequences from both the first and or the second special nucleic acids, toward the N-Terminal, through and including 35 or 47 amino acids, see Example 11 for the 47 example, beginning from the DTG site and including the nucleotides from that code for the previous 35 or 47 amino acids before the DTG site). The nucleic acid polynucleotide of claim 22, where the polynucleotide comprises identical to the same corresponding amino acids in SEQ. ID. NO. 3, that is, identical to the sequences in SEQ. ID. NO. 3 including the sequences from both the first and or the second special nucleic acids, toward the N-Terminal, through and including 35 or 47 amino acids, see Example 11 for the 47 example, beginning from the DTG site and including the nucleotides from that code for the previous 35 or 47 amino acids before the DTG site).

An isolated nucleic acid molecule comprising a polynucleotide, said polynucleotide encoding a Hu-Asp polypeptide and having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:

(a) a nucleotide sequence encoding a Hu-Asp polypeptide selected from the group consisting of Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b), wherein said Hu-Asp1, Hu-Asp2(a) and Hu-Asp2(b) polypeptides have the complete amino acid sequence of SEQ ID No. 2, SEQ ID No. 4, and SEQ ID No. 6, respectively; and

(b) a nucleotide sequence complementary to the nucleotide sequence of (a).

The nucleic acid molecule of claim 92, wherein said Hu-Asp polypeptide is Hu-Asp1, and said polynucleotide molecule of 1(a) comprises the nucleotide sequence of SEQ ID No. 1. The nucleic acid molecule of claim 92, wherein said Hu-Asp polypeptide is Hu-

Asp2(a), and said polynucleotide molecule of 1(a) comprises the nucleotide sequence of SEQ ID No. 4. The nucleic acid molecule of claim 92, wherein said Hu-Asp polypeptide is Hu-Asp2(b), and said polynucleotide molecule of 1(a) comprises the nucleotide sequence of SEQ ID No. 5. An isolated nucleic acid molecule comprising polynucleotide which hybridizes under stringent conditions to a polynucleotide having the nucleotide sequence in (a) or (b) of claim 92. A vector comprising the nucleic acid molecule of claim 96. The vector of claim 97, wherein said nucleic acid molecule is operably linked to a promoter for the expression of a Hu-Asp polypeptide. The vector of claim 98, wherein said Hu-Asp polypeptide is Hu-Asp1. The vector of claim 98, wherein said Hu-Asp polypeptide is Hu-Asp2(a). The vector of claim 98, wherein said Hu-Asp polypeptide is Hu-Asp2(b). A host cell comprising the vector of claim 98. A method of obtaining a Hu-Asp polypeptide comprising culturing the host cell of claim 102 and isolating said Hu-Asp polypeptide. An isolated Hu-Asp1 polypeptide comprising an amino acid sequence at least 95% identical to a sequence comprising the amino acid sequence of SEQ ID No. 2. An isolated Hu-Asp2(a) polypeptide comprising an amino acid sequence at least 95% identical to a sequence comprising the amino acid sequence of SEQ ID No. 4. An isolated Hu-Asp2(a) polypeptide comprising an amino acid sequence at least 95% identical to a sequence comprising the amino acid sequence of SEQ ID No. 8. An isolated antibody that binds specifically to the Hu-Asp polypeptide of any of claims 104-107.

Here we disclose numerous methods to assay the enzyme.

A method to identify a cell that can be used to screen for inhibitors of β secretase activity comprising:

(a) identifying a cell that expresses a protease capable of cleaving APP at the β secretase site, comprising:

i) collect the cells or the supernantent from the cells to be identified

ii) measure the production of a critical peptide, where the critical peptide is selected from the group consisting of either the APP C-terminal peptide or soluble APP,

iii) select the cells which produce the critical peptide.

The method of claim 108 where the cells are collected and the critical peptide is the APP C-terminal peptide created as a result of the β secretase cleavage. The method of claim 108 where the supernantent is collected and the critical peptide is soluble APP where the soluble APP has a C-terminal created by β secretase cleavage. The method of claim 108

where the cells contain any of the nucleic acids or polypeptides of claims 1-86 and where the cells are shown to cleave the β secretase site of any peptide having the following peptide structure, P2, P1, P1', P2', where P2 is K or N, where P1 is M or L, where P1' is D, where P2' is A. The method of claim 111 where P2 is K and P1 is M.. The method of claim 112 where P2 is N and P1 is L.

Any bacterial cell comprising any nucleic acids or peptides in claims 1-86 and 92-107. A bacterial cell of claim 114 where the bacteria is *E coli*. Any eukaryotic cell comprising any nucleic acids or polypeptides in claims 1-86 and 92-107.

Any insect cell comprising any of the nucleic acids or polypeptides in claims 1-86 and 92-107. A insect cell of claim 117 where the insect is sf9, or High 5. A insect cell of claim 100 where the insect cell is High 5. A mammalian cell comprising any of the nucleic acids or polypeptides in claims 1-86 and 92-107. A mammalian cell of claim 120 where the mammalian cell is selected from the group consisting of, human, rodent, lagomorph, and primate. A mammalian cell of claim 121 where the mammalian cell is selected from the group consisting of human cell. A mammalian cell of claim 122 where the human cell is selected from the group comprising HEK293, and IMR-32. A mammalian cell of claim 121 where the cell is a primate cell. A primate cell of claim 124 where the primate cell is a COS-7 cell. A mammalian cell of claim 121 where cell is selected from a rodent cells. A rodent cell of claim 126 selected from, CHO-K1, Neuro-2A, 3T3 cells. A yeast cell of claim 115. An avian cell of claim 115.

Any isoform of APP where the last two carboxy terminus amino acids of that isoform are both lysine residues. In written descrip. Define isoform is any APP polypeptide, including APP variants (including mutations), and APP fragments that exists in humans such as those desribed in US 5,766,846, col 7, lines 45-67, incorporated into this document by reference. The isoform of APP from claim 114, comprising the isoform known as APP695 modified so that its last two having two lysine residues as its last two carboxy terminus amino acids. The isoform of claim 130 comprising SEQ. ID. 16. The isoform variant of claim 130 comprising SEQ. ID. NO. 18, and 20. Any eukaryotic cell line, comprising nucleic acids or polypeptides of claim 130-132. Any cell line of claim 133 that is a mammaliam cell line (HEK293, Neuro2a, best - plus others. A method for identifying inhibitors of an enzyme that cleaves the beta secretase cleavabe site of APP comprising:

a) culturing cells in a culture medium under conditions in which the enzyme causes processing of APP and release of amyloid beta-peptide into the medium and causes the accumulation of CTF99 fragments of APP in cell lysates.

b) exposing the cultured cells to a test compound; and specifically determining whether the test compound inhibits the function of the enzyme by measuring the amount of amyloid beta-peptide released into the medium and or the amount of CTF99 fragments of APP in cell lysates;

c) identifying test compounds diminishing the amount of soluble amyloid beta peptide present in the culture medium and diminution of CTF99 fragments of APP in cell lysates as Asp2 inhibitors.

The method of claim 135 wherein the cultured cells are a human, rodent or insect cell line. The method of claim 136 wherein the human or rodent cell line exhibits β secretase activity in which processing of APP occurs with release of amyloid beta-peptide into the culture medium and accumulation of CTF99 in cell lysates. A method as in claim 137 wherein the human or rodent cell line treated with the antisense oligomers directed against the enzyme that exhibits β secretase activity, reduces release of soluble amyloid beta-peptide into the culture medium and accumulation of CTF99 in cell lysates. A method for the identification of an agent that decreases the activity of a Hu-Asp polypeptide selected from the group consisting of Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b), the method comprising:

a) determining the activity of said Hu-Asp polypeptide in the presence of a test agent and in the absence of a test agent; and

b) comparing the activity of said Hu-Asp polypeptide determined in the presence of said test agent to the activity of said Hu-Asp polypeptide determined in the absence of said test agent;

whereby a lower level of activity in the presence of said test agent than in the absence of said test agent indicates that said test agent has decreased the activity of said Hu-Asp polypeptide..The nucleic acids, peptides, proteins, vectors, cells and cell lines, and assays described herein.

The present invention provides isolated nucleic acid molecules comprising a polynucleotide that codes for a polypeptide selected from the group consisting of human aspartyl proteases. In particular, human aspartyl protease 1 (Hu-Asp1) and two alternative splice variants of human aspartyl protease 2 (Hu-Asp2), designated herein as Hu-Asp2(a) and

Hu-Asp2(b). As used herein, all references to "Hu-Asp" should be understood to refer to all of Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b). In addition, as used herein, all references to "Hu-Asp2" should be understood to refer to both Hu-Asp2(a) and Hu-Asp2(b). Hu-Asp1 is expressed most abundantly in pancreas and prostate tissues, while Hu-Asp2(a) and Hu-Asp2(b) are expressed most abundantly in pancreas and brain tissues. The invention also provides isolated Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b) polypeptides, as well as fragments thereof which exhibit aspartyl protease activity.

In a preferred embodiment, the nucleic acid molecules comprise a polynucleotide having a nucleotide sequence selected from the group consisting of residues 1-1554 of SEQ ID NO:1, encoding Hu-Asp1, residues 1-1503 of SEQ ID NO:3, encoding Hu-Asp2(a), and residues 1-1428 of SEQ ID NO:5, encoding Hu-Asp2(b). In another aspect, the invention provides an isolated nucleic acid molecule comprising a polynucleotide which hybridizes under stringent conditions to a polynucleotide encoding Hu-Asp1, Hu-Asp2(a), Hu-Asp-2(b), or fragments thereof. European patent application EP 0 848 062 discloses a polypeptide referred to as "Asp 1," that bears substantial homology to Hu-Asp1, while international application WO 98/22597 discloses a polypeptide referred to as "Asp 2," that bears substantial homology to Hu-Asp2(a).

The present invention also provides vectors comprising the isolated nucleic acid molecules of the invention, host cells into which such vectors have been introduced, and recombinant methods of obtaining a Hu-Asp1, Hu-Asp2(a), or Hu-Asp2(b) polypeptide comprising culturing the above-described host cell and isolating the relevant polypeptide.

In another aspect, the invention provides isolated Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b) polypeptides, as well as fragments thereof. In a preferred embodiment, the Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b) polypeptides have the amino acid sequence given in SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:6, respectively. The present invention also describes active forms of Hu-Asp2, methods for preparing such active forms, methods for preparing soluble forms, methods for measuring Hu-Asp2 activity, and substrates for Hu-Asp2 cleavage. The invention also describes antisense oligomers targeting the Hu-Asp1, Hu-Asp2(a) and Hu-Asp2(b) mRNA transcripts and the use of such antisense reagents to decrease such mRNA and consequently the production of the corresponding polypeptide. Isolated antibodies, both polyclonal and monoclonal, that binds specifically to any of the Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b) polypeptides of the invention are also provided.

The invention also provides a method for the identification of an agent that modulates the activity of any of Hu-Asp-1, Hu-Asp2(a), and Hu-Asp2(b). The inventions describes methods to test such agents in cell-free assays to which Hu-Asp2 polypeptide is added, as well as methods to test such agents in human or other mammalian cells in which Hu-Asp2 is present.

BRIEF DESCRIPTION OF THE SEQUENCE LISTINGS

Sequence ID No. 1—Human Asp-1, nucleotide sequence
Sequence ID No. 2—Human Asp-1, predicted amino acid sequence
Sequence ID No. 3—Human Asp-2(a), nucleotide sequence
Sequence ID No. 4—Human Asp-2(a), predicted amino acid sequence
Sequence ID No. 5—Human Asp-2(b), nucleotide sequence
Sequence ID No. 6—Human Asp-2(b), predicted amino acid sequence
Sequence ID No. 7—Murine Asp-2(a), nucleotide sequence
Sequence ID No. 8—Murine Asp-2(a), predicted amino acid sequence
Sequence ID No. 9—Human APP695, nucleotide sequence
Sequence ID No.10—Human APP695, predicted amino acid sequence
Sequence ID No.11—Human APP695-Sw, nucleotide sequence
Sequence ID No.12—Human APP695-Sw. predicted amino acid sequence
Sequence ID No.13—Human APP695-VF, nucleotide sequence
Sequence ID No.14—Human APP695-VF, predicted amino acid sequence
Sequence ID No.15—Human APP695-KK, nucleotide sequence
Sequence ID No.16—Human APP695-KK, predicted amino acid sequence
Sequence ID No.17—Human APP695-Sw-KK, nucleotide sequence
Sequence ID No.18—Human APP695-Sw-KK, predicted amino acid sequence
Sequence ID No.19—Human APP695-VF-KK, nucleotide sequence
Sequence ID No.20—Human APP695-VF-KK, predicted amino acid sequence
Sequence ID No.21—T7-Human-pro-Asp-2(a)ΔTM, nucleotide sequence
Sequence ID No.22—T7-Human-pro-Asp-2(a)ΔTM, amino acid sequence
Sequence ID No.23—T7-Caspase-Human-pro-Asp-2(a)ΔTM, nucleotide sequence
Sequence ID No.24—T7-Caspase-Human-pro-Asp-2(a)ΔTM, amino acid sequence
Sequence ID No.25—Human-pro-Asp-2(a)ΔTM (low GC), nucleotide sequence
Sequence ID No.26—Human-pro-Asp-2(a)ΔTM, (low GC), amino acid sequence
Sequence ID No.27—T7-Caspase-Caspase 8 cleavage-Human-pro-Asp-2(a)ΔTM, nucleotide sequence
Sequence ID No.28—T7-Caspase-Caspase 8 cleavage-Human-pro-Asp-2(a)ΔTM, amino acid sequence
Sequence ID No.29—Human Asp-2(a)ΔTM, nucleotide sequence
Sequence ID No.30—Human Asp-2(a)ΔTM, amino acid sequence
Sequence ID No.31—Human Asp-2(a)ΔTM$(His)_6$, nucleotide sequence
Sequence ID No.32—Human Asp-2(a)ΔTM$(His)_6$, amino acid sequence
Sequence ID No.s 33-46 are described below in the Detailed Description of the Invention.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1: Figure 1 shows the nucleotide (SEQ ID NO:1) and predicted amino acid sequence (SEQ ID NO:2) of human Asp1.

Figure 2: Figure 2 shows the nucleotide (SEQ ID NO:3) and predicted amino acid sequence (SEQ ID NO:4) of human Asp2(a).

Figure 3: Figure 3 shows the nucleotide (SEQ ID NO:5) and predicted amino acid sequence (SEQ ID NO:6) of human Asp2(b). The predicted transmembrane domain of Hu-Asp2(b) is enclosed in brackets.

Figure 4: Figure 4 shows the nucleotide (SEQ ID No. 7) and predicted amino acid sequence (SEQ ID No. 8) of murine Assp2(a)

Figure 5: Figure 5 shows the BestFit alignment of the predicted amino acid sequences of Hu-Asp2(a) and murine Asp2(a)

Figure 6: Figure 6 shows the nucleotide (SEQ ID No. 21) and predicted amino acid sequence (SEQ ID No. 22) of T7-Human-pro-Asp-2(a)ΔTM

Figure7: Figure 7 shows the nucleotide (SEQ ID No. 23) and predicted amino acid sequence (SEQ ID No. 24) of T7-caspase-Human-pro-Asp-2(a)ΔTM

Figure 8: Figure 8 shows the nucleotide (SEQ ID No. 25) and predicted amino acid sequence (SEQ ID No. 26) of Human-pro-Asp-2(a)ΔTM (low GC)

Figure 9: Western blot showing reduction of CTF99 production by HEK125.3 cells transfected with antisense oligomers targeting the Hu-Asp2 Mma

Figure 10: Western blot showing increase in CTF99 production in mouse Neuro-2a cells cotransfected with APP-KK with and without Hu-Asp2 only in those cells cotransfected with Hu-Asp2. A further increase in CTF99 production is seen in cells cotransfected with APP-Sw-KK with and without Hu-Asp2 only in those cells cotransfected with Hu-Asp2

Figure 11: Figure 11 shows the predicted amino acid sequence (SEQ ID No. 30) of Human-Asp2(a)ΔTM

Figure 12: Figure 11 shows the predicted amino acid sequence (SEQ ID No. 30) of Human-Asp2(a)ΔTM(His)$_6$

DETAILED DESCRIPTION OF THE INVENTION

A few definitions used in this invention follow, most definitions to be used are those that would be used by one ordinarily skilled in the art.

When the β amyloid peptide any peptide resulting from beta secretase cleavage of APP. This includes, peptides of 39, 40, 41, 42 and 43 amino acids, extending from the β-

secretase cleavage site to 39, 40, 41, 42 and 43 amino acids. β amyloid peptide also means sequences 1-6, SEQ. ID. NO. 1-6 of US 5,750,349, issued 12 May 1998 (incorporated into this document by reference). A β-secretase cleavage fragment disclosed here is called CTF-99, which extends from β-secretase cleavage site to the carboxy terminus of APP.

When an isoform of APP is discussed then what is meant is any APP polypeptide, including APP variants (including mutations), and APP fragments that exists in humans such as those desribed in US 5,766,846, col 7, lines 45-67, incorporated into this document by reference and see below.

The term "β-amyloid precursor protein" (APP) as used herein is defined as a polypeptide that is encoded by a gene of the same name localized in humans on the long arm of chromosome 21 and that includes "βAP – here "β-amyloid protein" see above, within its carboxyl third. APP is a glycosylated, single-membrane spanning protein expressed in a wide variety of cells in many mammaliam tissues. Examples of specific isotypes of APP which are currently known to exist in humans are the 695-amino acid polypeptide described by Kang et. al. (1987) Nature 325:733-736 which is designated as the "normal" APP; the 751-amino acid polypeptide described by Ponte et al. (1988) Nature 331:525-527 (1988) and Tanzi et al. (1988) Nature 331:528-530; and the 770-amino acid polypeptide described by Kitaguchi et. al. (1988) Nature 331:530-532. Examples of specific variants of APP include point mutation which can differ in both position and phenotype (for review of known variant mutation see Hardy (1992) Nature Genet. 1:233-234). All references cited here incorporated by reference. The term "APP fragments" as used herein refers to fragments of APP other than those which consist solely of βAP or βAP fragments. That is, APP fragments will include amino acid sequences of APP in addition to those which form intact 3AP or a fragment of βAP.

When the term "any amino acid" is used, the amino acids referred to are to be selected from the following, three letter and single letter abbreviations - which may also be used, are provided as follows:

Alanine, Ala, A; Arginine, Arg, R; Asparagine, Asn, N; Aspartic acid, Asp, D; Cystein, Cys, C; Glutamine, Gln, Q; lu;E-Glutamic Acid, Glu, E; Glycine, Gly, G; Histidine, His, H; Isoleucine, Ile, I; Leucine, Leu, L; Lysine, Lys, K; Methionine, Met, M; Phenylalanine, Phe, F; Proline, Pro, P; Serine, Ser, S; Threonine, Thr, T; Tryptophan, Trp, W; Tyrosine, Tyr, Y; Valine, Val, V; Aspartic acid or Asparagine, Asx, B; Glutamic acid or Glutamine, Glx, Z; Any amino acid, Xaa, X..

The present invention describes a method to scan gene databases for the simple active site motif characteristic of aspartyl proteases. Eukaryotic aspartyl proteases such as pepsin and renin possess a two-domain structure which folds to bring two aspartyl residues into proximity within the active site. These are embedded in the short tripeptide motif DTG, or more rarely, DSG. Most aspartyl proteases occur as proenzyme whose N-terminus must be cleaved for activation. The DTG or DSG active site motif appears at about residue 65-70 in the proenzyme (prorenin, pepsinogen), but at about residue 25-30 in the active enzyme after cleavage of the N-terminal prodomain. The limited length of the active site motif makes it difficult to search collections of short, expressed sequence tags (EST) for novel aspartyl proteases. EST sequences typically average 250 nucleotides or less, and so would encode 80-90 amino acid residues or less. That would be too short a sequence to span the two active site motifs. The preferred method is to scan databases of hypothetical or assembled protein coding sequences. The present invention describes a computer method to identify candidate aspartyl proteases in protein sequence databases. The method was used to identify seven candidate aspartyl protease sequences in the Caenorhabditis elegans genome. These sequences were then used to identify by homology search Hu-Asp1 and two alternative splice variants of Hu-Asp2, designated herein as Hu-Asp2(a) and Hu-Asp2(b).

In a major aspect of the invention disclosed here we provide new information about APP processing. Pathogeneic processing of the amyloid precursor protein (APP) via the Aβ pathway requires the sequential action of two proteases referred to as β-secretase and γ-secretase. Cleavage of APP by the β-secretase and γ-secretase generates the N-terminus and C-terminus of the Aβ peptide, respectively. Because over production of the Aβ peptide, particularly the $A\beta_{1-42}$, has been implicated in the initiation of Alzheimer's disease, inhibitors of either the β-secretase and/or the γ-secretase have potential in the treatment of Alzheimer's disease. Despite the importance of the β-secretase and γ-secretase in the pathogenic processing of APP, molecular definition of these enzymes has not been accomplished to date. That is, it was not known what enzymes were required for cleavage at either the β-secretase or the γ-secretase cleavage site. The sites themselves were known because APP was known and the $A\beta_{1-42}$, peptide was known, see US 5,766,846 and US 5,837,672, (incorporated by reference, with the exception to reference to "soluble" peptides). But what enzyme was involved in producing the $A\beta_{1-42}$, peptide was unknown.

The present invention involves the molecular definition of several novel human aspartyl proteases and one of these, referred to as Hu-Asp-2(a) and Hu-Asp2(b), has been characterized in detail. Previous forms of asp1 and asp 2 have been disclosed, see EP 0848062 A2 and EP 0855444A2, inventors David Powel et. al., assigned to Smith Kline Beecham Corp. (incorporated by reference). Herein are disclosed old and new forms of Hu-Asp 2. For the first time they are expressed in active form, their substrates are disclosed, and their specificity is disclosed. Prior to this disclosure cell or cell extracts were required to cleave the β-secretase site, now purified protein can be used in assays, also described here. Based on the results of (1) antisense knock out experiments, (2) transient transfection knock in experiments, and (3) biochemical experiments using purified recombinant Hu-Asp-2, we demonstrate that Hu-Asp-2 is the β-secretase involved in the processing of APP. Although the nucleotide and predicted amino acid sequence of Hu-Asp-2(a) has been reported, see above, see EP 0848062 A2 and EP 0855444A2, no functional characterization of the enzyme was disclosed. Here the authors characterize the Hu-Asp-2 enzyme and are able to explain why it is a critical and essential enzyme required in the formation of $A\beta_{1-42}$, peptide and possible a critical step in the development of AD.

In another embodiment the present invention also describes a novel splice variant of Hu-Asp2, referred to as Hu-Asp-2(b), that has never before been disclosed.

In another embodiment, the invention provides isolated nucleic acid molecules comprising a polynucleotide encoding a polypeptide selected from the group consisting of human aspartyl protease 1 (Hu-Asp1) and two alternative splice variants of human aspartyl protease 2 (Hu-Asp2), designated herein as Hu-Asp2(a) and Hu-Asp2(b). As used herein, all references to "Hu-Asp2" should be understood to refer to both Hu-Asp2(a) and Hu-Asp2(b). Hu-Asp1 is expressed most abundantly in pancreas and prostate tissues, while Hu-Asp2(a) and Hu-Asp2(b) are expressed most abundantly in pancreas and brain tissues. The invention also provides isolated Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b) polypeptides, as well as fragments thereof which exhibit aspartyl protease activity.

The predicted amino acid sequences of Hu-Asp1, Hu-Asp2(a) and Hu-Asp2(b) share significant homology with previously identified mammalian aspartyl proteases such as pepsinogen A, pepsinogen B, cathepsin D, cathepsin E, and renin. P.B.Szecs, *Scand. J. Clin. Lab. Invest.* 52:(Suppl. 210 5-22 (1992)). These enzymes are characterized by the presence of a duplicated DTG/DSG sequence motif. The Hu-Asp1 and HuAsp2 polypeptides disclosed

herein also exhibit extremely high homology with the ProSite consensus motif for aspartyl proteases extracted from the SwissProt database.

The nucleotide sequence given as residues 1-1554 of SEQ ID NO:1 corresponds to the nucleotide sequence encoding Hu-Asp1, the nucleotide sequence given as residues 1-1503 of SEQ ID NO:3 corresponds to the nucleotide sequence encoding Hu-Asp2(a), and the nucleotide sequence given as residues 1-1428 of SEQ ID NO:5 corresponds to the nucleotide sequence encoding Hu-Asp2(b). The isolation and sequencing of DNA encoding Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b) is described below in Examples 1 and 2.

As is described in Examples 1 and 2, automated sequencing methods were used to obtain the nucleotide sequence of Hu-Asp1, Hu-Asp2(a), and Hu-Asp-2(b). The Hu-Asp nucleotide sequences of the present invention were obtained for both DNA strands, and are believed to be 100% accurate. However, as is known in the art, nucleotide sequence obtained by such automated methods may contain some errors. Nucleotide sequences determined by automation are typically at least about 90%, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of a given nucleic acid molecule. The actual sequence may be more precisely determined using manual sequencing methods, which are well known in the art. An error in sequence which results in an insertion or deletion of one or more nucleotides may result in a frame shift in translation such that the predicted amino acid sequence will differ from that which would be predicted from the actual nucleotide sequence of the nucleic acid molecule, starting at the point of the mutation. The Hu-Asp DNA of the present invention includes cDNA, chemically synthesized DNA, DNA isolated by PCR, genomic DNA, and combinations thereof. Genomic Hu-Asp DNA may be obtained by screening a genomic library with the Hu-Asp2 cDNA described herein, using methods that are well known in the art, or with oligonucleotides chosen from the Hu-Asp2 sequence that will prime the polymerase chain reaction (PCR). RNA transcribed from Hu-Asp DNA is also encompassed by the present invention.

Due to the degeneracy of the genetic code, two DNA sequences may differ and yet encode identical amino acid sequences. The present invention thus provides isolated nucleic acid molecules having a polynucleotide sequence encoding any of the Hu-Asp polypeptides of the invention, wherein said polynucleotide sequence encodes a Hu-Asp polypeptide having the complete amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, or fragments thereof.

Also provided herein are purified Hu-Asp polypeptides, both recombinant and non-recombinant. Most importantly, methods to produce Hu-Asp2 polypeptides in active form are provided. These include production of Hu-Asp2 polypeptides and variants thereof in bacterial cells, insect cells, and mammalian cells, also in forms that allow secretion of the Hu-Asp2 polypeptide from bacterial, insect or mammalian cells into the culture medium, also methods to produce variants of Hu-Asp2 polypeptide incorporating amino acid tags that facilitate subsequent purification. In a preferred embodiment of the invention the Hu-Asp2 polypeptide is converted to a proteolytically active form either in transformed cells or after purification and cleavage by a second protease in a cell-free system, such active forms of the Hu-Asp2 polypeptide beginning with the N-terminal sequence TQHGIR or ETDEEP. Variants and derivatives, including fragments, of Hu-Asp proteins having the native amino acid sequences given in SEQ ID Nos: 2, 4, and 6 that retain any of the biological activities of Hu-Asp are also within the scope of the present invention. Of course, one of ordinary skill in the art will readily be able to determine whether a variant, derivative, or fragment of a Hu-Asp protein displays Hu-Asp activity by subjecting the variant, derivative, or fragment to a standard aspartyl protease assay. Fragments of Hu-Asp within the scope of this invention include those that contain the active site domain containing the amino acid sequence DTG, fragments that contain the active site domain amino acid sequence DSG, fragments containing both the DTG and DSG active site sequences, fragments in which the spacing of the DTG and DSG active site sequences has been lengthened, fragments in which the spacing has been shortened. Also within the scope of the invention are fragments of Hu-Asp in which the transmembrane domain has been removed to allow production of Hu-Asp2 in a soluble form. In another embodiment of the invention, the two halves of Hu-Asp2, each containing a single active site DTG or DSG sequence can be produced independently as recombinant polypeptides, then combined in solution where they reconstitute an active protease.

Hu-Asp variants may be obtained by mutation of native Hu-Asp-encoding nucleotide sequences, for example. A Hu-Asp variant, as referred to herein, is a polypeptide substantially homologous to a native Hu-Asp polypeptide but which has an amino acid sequence different from that of native Hu-Asp because of one or more deletions, insertions, or substitutions in the amino acid sequence. The variant amino acid or nucleotide sequence is preferably at least about 80% identical, more preferably at least about 90% identical, and most preferably at least about 95% identical, to a native Hu-Asp sequence. Thus, a variant nucleotide sequence which contains, for example, 5 point mutations for every one hundred

nucleotides, as compared to a native Hu-Asp gene, will be 95% identical to the native protein. The percentage of sequence identity, also termed homology, between a native and a variant Hu-Asp sequence may also be determined, for example, by comparing the two sequences using any of the computer programs commonly employed for this purpose, such as the Gap program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, Madison Wisconsin), which uses the algorithm of Smith and Waterman (*Adv. Appl. Math. 2:* 482-489 (1981)).

Alterations of the native amino acid sequence may be accomplished by any of a number of known techniques. For example, mutations may be introduced at particular locations by procedures well known to the skilled artisan, such as oligonucleotide-directed mutagenesis, which is described by Walder *et al.* (*Gene 42*:133 (1986)); Bauer *et al.* (*Gene 37*:73 (1985)); Craik (*BioTechniques*, January 1985, pp. 12-19); Smith *et al.* (*Genetic Engineering: Principles and Methods*, Plenum Press (1981)); and U.S. Patent Nos. 4,518,584 and 4,737,462.

Hu-Asp variants within the scope of the invention may comprise conservatively substituted sequences, meaning that one or more amino acid residues of a Hu-Asp polypeptide are replaced by different residues that do not alter the secondary and/or tertiary structure of the Hu-Asp polypeptide. Such substitutions may include the replacement of an amino acid by a residue having similar physicochemical properties, such as substituting one aliphatic residue (Ile, Val, Leu or Ala) for another, or substitution between basic residues Lys and Arg, acidic residues Glu and Asp, amide residues Gln and Asn, hydroxyl residues Ser and Tyr, or aromatic residues Phe and Tyr. Further information regarding making phenotypically silent amino acid exchanges may be found in Bowie *et al., Science 247*:1306-1310 (1990). Other Hu-Asp variants which might retain substantially the biological activities of Hu-Asp are those where amino acid substitutions have been made in areas outside functional regions of the protein.

In another aspect, the invention provides an isolated nucleic acid molecule comprising a polynucleotide which hybridizes under stringent conditions to a portion of the nucleic acid molecules described above, *e.g.*, to at least about 15 nucleotides, preferably to at least about 20 nucleotides, more preferably to at least about 30 nucleotides, and still more preferably to at least about from 30 to at least about 100 nucleotides, of one of the previously described nucleic acid molecules. Such portions of nucleic acid molecules having the described lengths refer to, *e.g.*, at least about 15 contiguous nucleotides of the reference nucleic acid molecule.

By stringent hybridization conditions is intended overnight incubation at about 42°C for about 2.5 hours in 6 X SSC/0.1% SDS, followed by washing of the filters in 1.0 X SSC at 65°C, 0.1% SDS.

Fragments of the Hu-Asp-encoding nucleic acid molecules described herein, as well as polynucleotides capable of hybridizing to such nucleic acid molecules may be used as a probe or as primers in a polymerase chain reaction (PCR). Such probes may be used, *e.g.*, to detect the presence of Hu-Asp nucleic acids in *in vitro* assays, as well as in Southern and northern blots. Cell types expressing Hu-Asp may also be identified by the use of such probes. Such procedures are well known, and the skilled artisan will be able to choose a probe of a length suitable to the particular application. For PCR, 5' and 3' primers corresponding to the termini of a desired Hu-Asp nucleic acid molecule are employed to isolate and amplify that sequence using conventional techniques.

Other useful fragments of the Hu-Asp nucleic acid molecules are antisense or sense oligonucleotides comprising a single-stranded nucleic acid sequence capable of binding to a target Hu-Asp mRNA (using a sense strand), or Hu-Asp DNA (using an antisense strand) sequence. In a preferred embodiment of the invention these Hu-Asp antisense oligonucleotides reduce Hu-Asp mRNA and consequent production of Hu-Asp polypeptides.

In another aspect, the invention includes Hu-Asp polypeptides with or without associated native pattern glycosylation. Both Hu-Asp1 and Hu-Asp2 have canonical acceptor sites for Asn-linked sugars, with Hu-Asp1 having two of such sites, and Hu-Asp2 having four. Hu-Asp expressed in yeast or mammalian expression systems (discussed below) may be similar to or significantly different from a native Hu-Asp polypeptide in molecular weight and glycosylation pattern. Expression of Hu-Asp in bacterial expression systems will provide non-glycosylated Hu-Asp.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. Hu-Asp polypeptides may be recovered and purified from tissues, cultured cells, or recombinant cell cultures by well-known methods, including ammonium sulfate or ethanol precipitation, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography, lectin chromatography, and high performance liquid chromatography (HPLC). In a preferred embodiment, an amino acid tag is added to the Hu-Asp polypeptide using genetic engineering techniques that are well known to practioners of the art which include addition of six histidine amino acid residues to allow

purification by binding to nickel immobilized on a suitable support, epitopes for polyclonal or monoclonal antibodies including but not limited to the T7 epitope, the myc epitope, and the V5a epitope, and fusion of Hu-Asp2 to suitable protein partners including but not limited to glutathione-S-transferase or maltose binding protien. In a preferred embodiment these additional amino acid sequences are added to the C-terminus of Hu-Asp but may be added to the N-terminus or at intervening positions within the Hu-Asp2 polypeptide.

The present invention also relates to vectors comprising the polynucleotide molecules of the invention, as well as host cell transformed with such vectors. Any of the polynucleotide molecules of the invention may be joined to a vector, which generally includes a selectable marker and an origin of replication, for propagation in a host. Because the invention also provides Hu-Asp polypeptides expressed from the polynucleotide molecules described above, vectors for the expression of Hu-Asp are preferred. The vectors include DNA encoding any of the Hu-Asp polypeptides described above or below, operably linked to suitable transcriptional or translational regulatory sequences, such as those derived from a mammalian, microbial, viral, or insect gene. Examples of regulatory sequences include transcriptional promoters, operators, or enhancers, mRNA ribosomal binding sites, and appropriate sequences which control transcription and translation. Nucleotide sequences are operably linked when the regulatory sequence functionally relates to the DNA encoding Hu-Asp. Thus, a promoter nucleotide sequence is operably linked to a Hu-Asp DNA sequence if the promoter nucleotide sequence directs the transcription of the Hu-Asp sequence.

Selection of suitable vectors to be used for the cloning of polynucleotide molecules encoding Hu-Asp, or for the expression of Hu-Asp polypeptides, will of course depend upon the host cell in which the vector will be transformed, and, where applicable, the host cell from which the Hu-Asp polypeptide is to be expressed. Suitable host cells for expression of Hu-Asp polypeptides include prokaryotes, yeast, and higher eukaryotic cells, each of which is discussed below.

The Hu-Asp polypeptides to be expressed in such host cells may also be fusion proteins which include regions from heterologous proteins. Such regions may be included to allow, *e.g.*, secretion, improved stability, or facilitated purification of the polypeptide. For example, a sequence encoding an appropriate signal peptide can be incorporated into expression vectors. A DNA sequence for a signal peptide (secretory leader) may be fused in-frame to the Hu-Asp sequence so that Hu-Asp is translated as a fusion protein comprising the signal peptide. A signal peptide that is functional in the intended host cell promotes

extracellular secretion of the Hu-Asp polypeptide. Preferably, the signal sequence will be cleaved from the Hu-Asp polypeptide upon secretion of Hu-Asp from the cell. Non-limiting examples of signal sequences that can be used in practicing the invention include the yeast I-factor and the honeybee melatin leader in sf9 insect cells.

In a preferred embodiment, the Hu-Asp polypeptide will be a fusion protein which includes a heterologous region used to facilitate purification of the polypeptide. Many of the available peptides used for such a function allow selective binding of the fusion protein to a binding partner. For example, the Hu-Asp polypeptide may be modified to comprise a peptide to form a fusion protein which specifically binds to a binding partner, or peptide tag. Non-limiting examples of such peptide tags include the 6-His tag, thioredoxin tag, hemaglutinin tag, GST tag, and OmpA signal sequence tag. As will be understood by one of skill in the art, the binding partner which recognizes and binds to the peptide may be any molecule or compound including metal ions (*e.g.*, metal affinity columns), antibodies, or fragments thereof, and any protein or peptide which binds the peptide, such as the FLAG tag.

Suitable host cells for expression of Hu-Asp polypeptides includes prokaryotes, yeast, and higher eukaryotic cells. Suitable prokaryotic hosts to be used for the expression of Hu-Asp include bacteria of the genera *Escherichia, Bacillus, and Salmonella*, as well as members of the genera *Pseudomonas, Streptomyces*, and *Staphylococcus*. For expression in, *e.g., E. coli*, a Hu-Asp polypeptide may include an N-terminal methionine residue to facilitate expression of the recombinant polypeptide in a prokaryotic host. The N-terminal Met may optionally then be cleaved from the expressed Hu-Asp polypeptide. Other N-terminal amino acid residues can be added to the Hu-Asp polypeptide to facilitate expression in Escherichia coli including but not limited to the T7 leader sequence, the T7-caspase 8 leader sequence, as well as others leaders including tags for purification such as the 6-His tag (Example 9). Hu-Asp polypeptides expressed in E. coli may be shortened by removal of the cytoplasmic tail, the transmembrane domain, or the membrane proximal region. Hu-Asp polypeptides expressed in E. coli may be obtained in either a soluble form or as an insoluble form which may or may not be present as an inclusion body. The insoluble polypeptide may be rendered soluble by guanidine HCl, urea or other protein denaturants, then refolded into a soluble form before or after purification by dilution or dialysis into a suitable aqueous buffer. If the inactive proform of the Hu-Asp was produced using recombinant methods, it may be rendered active by cleaving off the prosegment with a second suitable protease such as human immunodeficiency virus protease.

Expression vectors for use in prokaryotic hosts generally comprises one or more phenotypic selectable marker genes. Such genes generally encode, *e.g.*, a protein that confers antibiotic resistance or that supplies an auxotrophic requirement. A wide variety of such vectors are readily available from commercial sources. Examples include pSPORT vectors, pGEM vectors (Promega), pPROEX vectors (LTI, Bethesda, MD), Bluescript vectors (Stratagene), pET vectors (Novagen) and pQE vectors (Qiagen).

Hu-Asp may also be expressed in yeast host cells from genera including *Saccharomyces, Pichia,* and *Kluveromyces.* Preferred yeast hosts are *S. cerevisiae* and *P. pastoris.* Yeast vectors will often contain an origin of replication sequence from a 2T yeast plasmid, an autonomously replicating sequence (ARS), a promoter region, sequences for polyadenylation, sequences for transcription termination, and a selectable marker gene. Vectors replicable in both yeast and *E. coli* (termed shuttle vectors) may also be used. In addition to the above-mentioned features of yeast vectors, a shuttle vector will also include sequences for replication and selection in *E. coli.* Direct secretion of Hu-Asp polypeptides expressed in yeast hosts may be accomplished by the inclusion of nucleotide sequence encoding the yeast I-factor leader sequence at the 5' end of the Hu-Asp-encoding nucleotide sequence.

Insect host cell culture systems may also be used for the expression of Hu-Asp polypeptides. In a preferred embodiment, the Hu-Asp polypeptides of the invention are expressed using an insect cell expression system (*see* Example 10). Additionally, a baculovirus expression system can be used for expression in insect cells as reviewed by Luckow and Summers, *Bio/Technology* 6:47 (1988).

In another preferred embodiment, the Hu-Asp polypeptide is expressed in mammalian host cells. Non-limiting examples of suitable mammalian cell lines include the COS-7 line of monkey kidney cells (Gluzman *et al., Cell* 23:175 (1981)), human embyonic kidney cell line 293, and Chinese hamster ovary (CHO) cells. Preferably, Chinese hamster ovary (CHO) cells are used for expression of Hu-Asp proteins (Example 11).

The choice of a suitable expression vector for expression of the Hu-Asp polypeptides of the invention will of course depend upon the specific mammalian host cell to be used, and is within the skill of the ordinary artisan. Examples of suitable expression vectors include pcDNA3 (Invitrogen) and pSVL (Pharmacia Biotech). A preferred vector for expression of Hu-Asp polypeptides is pcDNA3.1-Hygro (Invitrogen). Expression vectors for use in mammalian host cells may include transcriptional and translational control sequences derived

from viral genomes. Commonly used promoter sequences and enhancer sequences which may be used in the present invention include, but are not limited to, those derived from human cytomegalovirus (CMV), Adenovirus 2, Polyoma virus, and Simian virus 40 (SV40). Methods for the construction of mammalian expression vectors are disclosed, for example, in Okayama and Berg (*Mol. Cell. Biol. 3*:280 (1983)); Cosman *et al.* (*Mol. Immunol. 23*:935 (1986)); Cosman *et al.* (*Nature 312*:768 (1984)); EP-A-0367566; and WO 91/18982.

The polypeptides of the present invention may also be used to raise polyclonal and monoclonal antibodies, which are useful in diagnostic assays for detecting Hu-Asp polypeptide expression. Such antibodies may be prepared by conventional techniques. See, for example, *Antibodies: A Laboratory Manual*, Harlow and Land (eds.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., (1988); *Monoclonal Antibodies, Hybridomas: A New Dimension in Biological Analyses*, Kennet *et al.* (eds.), Plenum Press, New York (1980). Synthetic peptides comprising portions of Hu-Asp containing 5 to 20 amino acids may also be used for the production of polyclonal or monoclonal antibodies after linkage to a suitable carrier protein including but not limited to keyhole limpet hemacyanin (KLH), chicken ovalbumin, or bovine serum albumin using various cross-linking reagents including carbodimides, glutaraldehyde, or if the peptide contains a cysteine, N-methylmaleimide. A preferred peptide for immunization when conjugated to KLH contains the C-terminus of Hu_Asp1 or Hu-Asp2 comprising QRRPRDPEVVNDESSLVRHRWK or LRQQHDDFADDISLLK, respectively.

The Hu-Asp nucleic acid molecules of the present invention are also valuable for chromosome identification, as they can hybridize with a specific location on a human chromosome. Hu-Asp1 has been localized to chromosome 21, while Hu-Asp2 has been localized to chromosome 11q23.3-24.1. There is a current need for identifying particular sites on the chromosome, as few chromosome marking reagents based on actual sequence data (repeat polymorphisms) are presently available for marking chromosomal location. Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. The relationship between genes and diseases that have been mapped to the same chromosomal region can then be identified through linkage analysis, wherein the coinheritance of physically adjacent genes is determined. Whether a gene appearing to be related to a particular disease is in fact the cause of the disease can then be determined by comparing the nucleic acid sequence between affected and unaffected individuals.

In another embodiment, the invention relates to a method of assaying Hu-Asp function, specifically Hu-Asp2 function which involves incubating in solution the Hu-Asp polypeptide with a suitable substrate including but not limited to a synthetic peptide containing the β-secretase cleavage site of APP, preferably one containing the mutation found in a Swedish kindred with inherited AD in which KM is changed to NL, such peptide comprising the sequence SEVNLDAEFR in an acidic buffering solution, preferably an acidic buffering solution of pH5.5 (see Example 12) using cleavage of the peptide monitored by high performance liquid chromatography as a measure of Hu-Asp proteolytic activity. Preferred assays for proteolytic activity utilize internally quenched peptide assay substrates. Such suitable substrates include peptides which have attached a paired flurophore and quencher including but not limited to coumarin and dinitrophenol, respectively, such that cleavage of the peptide by the Hu-Asp results in increased fluorescence due to physical separation of the flurophore and quencher. Preferred colorimetric assays of Hu-Asp proteolytic activity utilize other suitable substrates that include the P2 and P1 amino acids comprising the recognition site for cleavage linked to o-nitrophenol through an amide linkage, such that cleavage by the Hu-Asp results in an increase in optical density after altering the assay buffer to alkaline pH.

In another embodiment, the invention relates to a method for the identification of an agent that increases the activity of a Hu-Asp polypeptide selected from the group consisting of Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b), the method comprising

(a) determining the activity of said Hu-Asp polypeptide in the presence of a test agent and in the absence of a test agent; and

(b) comparing the activity of said Hu-Asp polypeptide determined in the presence of said test agent to the activity of said Hu-Asp polypeptide determined in the absence of said test agent;

whereby a higher level of activity in the presence of said test agent than in the absence of said test agent indicates that said test agent has increased the activity of said Hu-Asp polypeptide. Such tests can be performed with Hu-Asp polypeptide in a cell free system and with cultured cells that express Hu-Asp as well as variants or isoforms thereof.

In another embodiment, the invention relates to a method for the identification of an agent that decreases the activity of a Hu-Asp polypeptide selected from the group consisting of Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b), the method comprising

(a) determining the activity of said Hu-Asp polypeptide in the presence of a test agent and in the absence of a test agent; and

(b) comparing the activity of said Hu-Asp polypeptide determined in the presence of said test agent to the activity of said Hu-Asp polypeptide determined in the absence of said test agent;

whereby a lower level of activity in the presence of said test agent than in the absence of said test agent indicates that said test agent has decreased the activity of said Hu-Asp polypeptide. Such tests can be performed with Hu-Asp polypeptide in a cell free system and with cultured cells that express Hu-Asp as well as variants or isoforms thereof.

In another embodiment, the invention relates to a novel cell line (HEK125.3 cells) for measuring processing of amyloid β peptide (Aβ) from the amyloid protein precursor (APP). The cells are stable transformants of human embryonic kidney 293 cells (HEK293) with a bicistronic vector derived from pIRES-EGFP (Clontech) containing a modified human APP cDNA, an internal ribosome entry site and an enhanced green fluorescent protein (EGFP) cDNA in the second cistron. The APP cDNA was modified by adding two lysine codons to the carboxyl terminus of the APP coding sequence. This increases processing of Aβ peptide from human APP by 2-4 fold. This level of Aβ peptide processing is 60 fold higher than is seen in nontransformed HEK293 cells. HEK125.3 cells will be useful for assays of compounds that inhibit Aβ peptide processing. This invention also includes addition of two lysine residues to the C-terminus of other APP isoforms including the 751 and 770 amino acid isoforms, to isoforms of APP having mutations found in human AD including the Swedish KM→NL and V717→F mutations, to C-terminal fragments of APP, such as those beginning with the β-secretase cleavage site, to C-terminal fragments of APP containing the β-secretase cleavage site which have been operably linked to an N-terminal signal peptide for membrane insertion and secretion, and to C-terminal fragments of APP which have been operably linked to an N-terminal signal peptide for membrane insertion and secretion and a reporter sequence including but not limited to green fluorescent protein or alkaline phosphatase, such that β-secretase cleavage releases the reporter protein from the surface of cells expressing the polypeptide.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

EXAMPLES

Example 1: Development of a Search Algorithm Useful for the Identification of Aspartyl Proteases, and Identification of C. elegans Aspartyl Protease Genes in Wormpep 12:

Materials and Methods:

Classical aspartyl proteases such as pepsin and renin possess a two-domain structure which folds to bring two aspartyl residues into proximity within the active site. These are embedded in the short tripeptide motif DTG, or more rarely, DSG. The DTG or DSG active site motif appears at about residue 25-30 in the enzyme, but at about 65-70 in the proenzyme (prorenin, pepsinogen). This motif appears again about 150-200 residues downstream. The proenzyme is activated by cleavage of the N-terminal prodomain. This pattern exemplifies the double domain structure of the modern day aspartyl enzymes which apparently arose by gene duplication and divergence. Thus;

$$NH_2\text{---------}X\text{-----}D^{25}TG\text{------------}Y\text{------}D^{Y+25}TG\text{-------------}C$$

where X denotes the beginning of the enzyme, following the N-terminal prodomain, and Y denotes the center of the molecule where the gene repeat begins again.

In the case of the retroviral enzymes such as the HIV protease, they represent only a half of the two-domain structures of well-known enzymes like pepsin, cathepsin D, renin, etc. They have no prosegment, but are carved out of a polyprotein precursor containing the *gag* and *pol* proteins of the virus. They can be represented by:

$$NH_2\text{--------}D^{25}TG\text{------------------}C100$$

This "monomer" only has about 100 aa, so is extremely parsimonious as compared to the other aspartyl protease "dimers" which have of the order of 330 or so aa, not counting the N-terminal prodomain.

The limited length of the eukaryotic aspartyl protease active site motif makes it difficult to search EST collections for novel sequences. EST sequences typically average 250 nucleotides, and so in this case would be unlikely to span both aspartyl protease active site motifs. Instead, we turned to the *C. elegans* genome. The *C. elegans* genome is estimated to contain around 13,000 genes. Of these, roughly 12,000 have been sequenced and the corresponding hypothetical open reading frame (ORF) has been placed in the database Wormpep12. We used this database as the basis for a whole genome scan of a higher eukaryote for novel aspartyl proteases, using an algorithm that we developed

specifically for this purpose. The following AWK script for locating proteins containing two DTG or DSG motifs was used for the search, which was repeated four times to recover all pairwise combinations of the aspartyl motif.

```
BEGIN{RS=">"}                /* defines ">" as record separator for FASTA format */
{
pos = index($0,"DTG")        /*finds "DTG" in record*/
if (pos>0) {
        rest = substr($0,pos+3)         /*get rest of record after first DTG*/
        pos2 = index(rest,"DTG")     /*find second DTG*/
        if (pos2>0) printf ("%s%s\n",">",$0)}          /*report hits*/
}
}
```

The AWK script shown above was used to search Wormpep12, which was downloaded from ftp.sanger.ac.uk/pub/databases/wormpep, for sequence entries containing at least two DTG or DSG motifs. Using AWK limited each record to 3000 characters or less. Thus, 35 or so larger records were eliminated manually from Wormpep12 as in any case these were unlikely to encode aspartyl proteases.

Results and Discussion:

The Wormpep 12 database contains 12,178 entries, although some of these (<10%) represent alternatively spliced transcripts from the same gene. Estimates of the number of genes encoded in the *C. elegans* genome is on the order of 13,000 genes, so Wormpep12 may be estimated to cover greater than 90% of the *C. elegans* genome.

Eukaryotic aspartyl proteases contain a two-domain structure, probably arising from ancestral gene duplication. Each domain contains the active site motif D(S/T)G located from 20-25 amino acid residues into each domain. The retroviral (e.g., HIV protease) or retrotransposon proteases are homodimers of subunits which are homologous to a single eukaryotic aspartyl protease domain. An AWK script was used to search the Wormpep12 database for proteins in which the D(S/T)G motif occurred at least twice. This identified >60 proteins with two DTG or DSG motifs. Visual inspection was used to select proteins in which the position of the aspartyl domains was suggestive of a two-domain structure meeting the criteria described above.

In addition, the PROSITE eukaryotic and viral aspartyl protease active site pattern PS00141 was used to search Wormpep12 for candidate aspartyl proteases. (Bairoch A., Bucher P., Hofmann K., The PROSITE database: its status in 1997, *Nucleic Acids Res.* *24*:217-221(1997)). This generated an overlapping set of Wormpep12 sequences. Of these,

seven sequences contained two DTG or DSG motifs and the PROSITE aspartyl protease active site pattern. Of these seven, three were found in the same cosmid clone (F21F8.3, F21F8.4, and F21F8.7) suggesting that they represent a family of proteins that arose by ancestral gene duplication. Two other ORFs with extensive homology to F21F8.3, F21F8.4 and F21F8.7 are present in the same gene cluster (F21F8.2 and F21F8.6), however, these contain only a single DTG motif. Exhaustive BLAST searches with these seven sequences against Wormpep12 failed to reveal additional candidate aspartyl proteases in the *C. elegans* genome containing two repeats of the DTG or DSG motif.

BLASTX search with each *C. elegans* sequence against SWISS-PROT, GenPep and TREMBL revealed that R12H7.2 was the closest worm homologue to the known mammalian aspartyl proteases, and that T18H9.2 was somewhat more distantly related, while CEASP1, F21F8.3, F21F8.4, and F21F8.7 formed a subcluster which had the least sequence homology to the mammalian sequences.

Discussion:

APP, the presenilins, and p35, the activator of cdk5, all undergo intracellular proteolytic processing at sites which conform to the substrate specificity of the HIV protease. Dysregulation of a cellular aspartyl protease with the same substrate specificity, might therefore provide a unifying mechanism for causation of the plaque and tangle pathologies in AD. Therefore, we sought to identify novel human aspartyl proteases. A whole genome scan in *C. elegans* identified seven open reading frames that adhere to the aspartyl protease profile that we had identified. These seven aspartyl proteases probably comprise the complete complement of such proteases in a simple, multicellular eukaryote. These include four closely related aspartyl proteases unique to *C. elegans* which probably arose by duplication of an ancestral gene. The other three candidate aspartyl proteases (T18H9.2, R12H7.2 and C11D2.2) were found to have homology to mammalian gene sequences.

Example 2: Identification of Novel Human Aspartyl Proteases Using Database Mining by Genome Bridging

Materials and Methods:

Computer-assisted analysis of EST databases, cDNA, and predicted polypeptide sequences:

Exhaustive homology searches of EST databases with the CEASP1, F21F8.3, F21F8.4, and F21F8.7 sequences failed to reveal any novel mammalian homologues. TBLASTN searches with R12H7.2 showed homology to cathepsin D, cathepsin E, pepsinogen A, pepsinogen C and renin, particularly around the DTG motif within the active site, but also failed to identify any additional novel mammalian aspartyl proteases. This indicates that the *C. elegans* genome probably contains only a single lysosomal aspartyl protease which in mammals is represented by a gene family that arose through duplication and consequent modification of an ancestral gene.

TBLASTN searches with T18H9.2, the remaining *C. elegans* sequence, identified several ESTs which assembled into a contig encoding a novel human aspartyl protease (Hu-ASP1). As is described above in Example 1, BLASTX search with the Hu-ASP1 contig against SWISS-PROT revealed that the active site motifs in the sequence aligned with the active sites of other aspartyl proteases. Exhaustive, repetitive rounds of BLASTN searches against LifeSeq, LifeSeqFL, and the public EST collections identified 102 EST from multiple cDNA libraries that assembled into a single contig. The 51 sequences in this contig found in public EST collections also have been assembled into a single contig (THC213329) by The Institute for Genome Research (TIGR). The TIGR annotation indicates that they failed to find any hits in the database for the contig. Note that the TIGR contig is the reverse complement of the LifeSeq contig that we assembled. BLASTN search of Hu-ASP1 against the rat and mouse EST sequences in ZooSeq revealed one homologous EST in each database (Incyte clone 700311523 and IMAGE clone 313341, GenBank accession number W10530, respectively).

TBLASTN searches with the assembled DNA sequence for Hu-ASP1 against both LifeSeqFL and the public EST databases identified a second, related human sequence (Hu-Asp2) represented by a single EST (2696295). Translation of this partial cDNA sequence reveals a single DTG motif which has homology to the active site motif of a bovine aspartyl protease, NM1.

BLAST searches, contig assemblies and multiple sequence alignments were performed using the bioinformatics tools provided with the LifeSeq, LifeSeqFL and LifeSeq Assembled databases from Incyte. Predicted protein motifs were identified using either the ProSite dictionary (Motifs in GCG 9) or the Pfam database.

Full-length cDNA cloning of Hu-Asp1

The open reading frame of *C. elegans* gene T18H9.2CE was used to query Incyte LifeSeq and LifeSeq-FL databases and a single electronic assembly referred to as 1863920CE1 was detected. The 5' most cDNA clone in this contig, 1863920, was obtained from Incyte and completely sequenced on both strands. Translation of the open reading frame contained within clone 1863920 revealed the presence of the duplicated aspartyl protease active site motif (DTG/DSG) but the 5' end was incomplete. The remainder of the Hu-Asp1 coding sequence was determined by 5' Marathon RACE analysis using a human placenta Marathon ready cDNA template (Clonetech). A 3'-antisense oligonucleotide primer specific for the 5' end of clone 1863920 was paired with the 5'-sense primer specific for the Marathon ready cDNA synthetic adaptor in the PCR. Specific PCR products were directly sequenced by cycle sequencing and the resulting sequence assembled with the sequence of clone 1863920 to yield the complete coding sequence of Hu-Asp-1 (SEQ ID No. 1).

Several interesting features are present in the primary amino acid sequence of Hu-Asp1 (Figure 1, SEQ ID No. 2). The sequence contains a signal peptide (residues 1-20 in SEQ ID No. 2), a pro-segment, and a catalytic domain containing two copies of the aspartyl protease active site motif (DTG/DSG). The spacing between the first and second active site motifs is about 200 residues which should correspond to the expected size of a single, eukaryotic aspartyl protease domain. More interestingly, the sequence contains a predicted transmembrane domain (residues 469-492 in SEQ ID No.2) near its C-terminus which suggests that the protease is anchored in the membrane. This feature is not found in any other aspartyl protease.

Cloning of a full-length Hu-Asp-2 cDNAs:

As is described above in Example 1, genome wide scan of the *Caenorhabditis elegans* database WormPep12 for putative aspartyl proteases and subsequent mining of human EST databases revealed a human ortholog to the *C. elegans* gene T18H9.2 referred to as Hu-Asp1. The assembled contig for Hu-Asp1 was used to query for human paralogs using the BLAST search tool in human EST databases and a single significant match

(2696295CE1) with approximately 60% shared identity was found in the LifeSeq FL database. Similar queries of either gb105PubEST or the family of human databases available from TIGR did not identify similar EST clones. cDNA clone 2696295, identified by single pass sequence analysis from a human uterus cDNA library, was obtained from Incyte and completely sequence on both strands. This clone contained an incomplete 1266 bp open-reading frame that encoded a 422 amino acid polypeptide but lacked an initiator ATG on the 5' end. Inspection of the predicted sequence revealed the presence of the duplicated aspartyl protease active site motif DTG/DSG, separated by 194 amino acid residues. Subsequent queries of later releases of the LifeSeq EST database identified an additional ESTs, sequenced from a human astrocyte cDNA library (4386993), that appeared to contain additional 5' sequence relative to clone 2696295. Clone 4386993 was obtained from Incyte and completely sequenced on both strands. Comparative analysis of clone 4386993 and clone 2696295 confirmed that clone 4386993 extended the open-reading frame by 31 amino acid residues including two in-frame translation initiation codons. Despite the presence of the two in-frame ATGs, no in-frame stop codon was observed upstream of the ATG indicating that the 4386993 may not be full-length. Furthermore, alignment of the sequences of clones 2696295 and 4386993 revealed a 75 base pair insertion in clone 2696295 relative to clone 4386993 that results in the insertion of 25 additional amino acid residues in 2696295. The remainder of the Hu-Asp2 coding sequence was determined by 5' Marathon RACE analysis using a human hippocampus Marathon ready cDNA template (Clonetech). A 3'-antisense oligonucleotide primer specific for the shared 5'-region of clones 2696295 and 4386993 was paired with the 5'-sense primer specific for the Marathon ready cDNA synthetic adaptor in the PCR. Specific PCR products were directly sequenced by cycle sequencing and the resulting sequence assembled with the sequence of clones 2696295 and 4386993 to yield the complete coding sequence of Hu-Asp2(a) (SEQ ID No. 3) and Hu-Asp2(b) (SEQ ID No. 5), respectively.

Several interesting features are present in the primary amino acid sequence of Hu-Asp2(a) (Figure 2 and SEQ ID No. 4) and Hu-Asp-2(b) (Figure 3, SEQ ID No. 6). Both sequences contain a signal peptide (residues 1-21 in SEQ ID No. 4 and SEQ ID No. 6), a pro-segment, and a catalytic domain containing two copies of the aspartyl protease active site motif (DTG/DSG). The spacing between the first and second active site motifs is variable due to the 25 amino acid residue deletion in Hu-Asp-2(b) and consists of 168-*versus*-194 amino acid residues, for Hu-Asp2(b) and Hu-Asp-2(a), respectively. More

interestingly, both sequences contains a predicted transmembrane domain (residues 455-477 in SEQ ID No.4 and 430-452 in SEQ ID No. 6) near their C-termini which indicates that the protease is anchored in the membrane. This feature is not found in any other aspartyl protease except Hu-Asp1.

Example 3. Molecular cloning of mouse Asp2 cDNA and genomic DNA.
Cloning and characterization of murine Asp2 cDNA—The murine ortholog of Hu_Asp2 was cloned using a combination of cDNA library screening, PCR, and genomic cloning. Approximately 500,000 independent clones from a mouse brain cDNA library were screened using a ^{32}P-labeled coding sequence probe prepared from Hu_Asp2. Replicate positives were subjected to DNA sequence analysis and the longest cDNA contained the entire 3' untranslated region and 47 amino acids in the coding region. PCR amplification of the same mouse brain cDNA library with an antisense oligonucleotide primer specific for the 5'-most cDNA sequence determined above and a sense primer specific for the 5' region of human Asp2 sequence followed by DNA sequence analysis gave an additional 980 bp of the coding sequence. The remainder of the 5' sequence of murine Asp-2 was derived from genomic sequence (see below).

Isolation and sequence analysis of the murine Asp-2 gene—A murine EST sequence encoding a portion of the murine Asp2 cDNA was identified in the GenBank EST database using the BLAST search tool and the Hu-Asp2 coding sequence as the query. Clone g3160898 displayed 88% shared identity to the human sequence over 352 bp. Oligonucleotide primer pairs specific for this region of murine Asp2 were then synthesized and used to amplify regions of the murine gene. Murine genomic DNA, derived from strain 129/SvJ, was amplified in the PCR (25 cycles) using various primer sets specific for murine Asp2 and the products analyzed by agarose gel electrophoresis. The primer set Zoo-1 and Zoo-4 amplified a 750 bp fragment that contained approximately 600 bp of intron sequence based on comparison to the known cDNA sequence. This primer set was then used to

screen a murine BAC library by PCR, a single genomic clone was isolated and this cloned was confirmed contain the murine Asp2 gene by DNA sequence analysis. Shotgun DNA sequencing of this Asp2 genomic clone and comparison to the cDNA sequences of both Hu_Asp2 and the partial murine cDNA sequences defined the full-length sequence of murine Asp2 (SEQ ID No. 7). The predicted amino acid sequence of murine Asp2 (SEQ ID No. 8) showed 96.4% shared identity (GCG BestFit algorithm) with 18/501 amino acid residue substitutions compared to the human sequence (Figure 4).

Example 4: Tissue Distribution of Expression of Hu-Asp2 Transcripts:

Materials and Methods:

The tissue distribution of expression of Hu-Asp-2 was determined using multiple tissue Northern blots obtained from Clonetech (Palo Alto, CA). Incyte clone 2696295 in the vector pINCY was digested to completion with *EcoRI/NotI* and the 1.8 kb cDNA insert purified by preparative agarose gel electrophoresis. This fragment was radiolabeled to a specific activity > 1 X 10^9 dpm/µg by random priming in the presence of [α-^{32}P-dATP] (>3000 Ci/mmol, Amersham, Arlington Heights, IL) and Klenow fragment of DNA polymerase I. Nylon filters containing denatured, size fractionated poly A^+ RNAs isolated from different human tissues were hybridized with 2 x 10^6 dpm/ml probe in ExpressHyb buffer (Clonetech, Palo Alto, CA) for 1 hour at 68 °C and washed as recommended by the manufacture. Hybridization signals were visualized by autoradiography using BioMax XR film (Kodak, Rochester, NY) with intensifying screens at -80 °C.

Results and Discussion:

Limited information on the tissue distribution of expression of Hu-Asp-2 transcripts was obtained from database analysis due to the relatively small number of ESTs detected using the methods described above (< 5). In an effort to gain further information on the expression of the Hu-Asp2 gene, Northern analysis was employed to determine both the size(s) and abundance of Hu-Asp2 transcripts. PolyA^+ RNAs isolated from a series of peripheral tissues and brain regions were displayed on a solid support following separation under denaturing conditions and Hu-Asp2 transcripts were visualized by high stringency hybridization to radiolabeled insert from clone 2696295. The 2696295 cDNA probe visualized a constellation of transcripts that migrated with apparent sizes of 3.0kb, 4.4 kb and 8.0 kb with the latter two transcript being the most abundant.

Across the tissues surveyed, Hu-Asp2 transcripts were most abundant in pancreas and brain with lower but detectable levels observed in all other tissues examined except thymus and PBLs. Given the relative abundance of Hu-Asp2 transcripts in brain, the regional expression in brain regions was also established. A similar constellation of transcript sizes were detected in all brain regions examined [cerebellum, cerebral cortex, occipital pole, frontal lobe, temporal lobe and putamen] with the highest abundance in the medulla and spinal cord.

Example 5: Northern Blot Detection of HuAsp-1 and HuAsp-2 Transcripts in Human Cell Lines:

A variety of human cell lines were tested for their ability to produce Hu-Asp1 and Asp2 mRNA. Human embryonic kidney (HEK-293) cells, African green monkey (Cos-7) cells, Chinese hamster ovary (CHO) cells, HELA cells, and the neuroblastoma cell line IMR-32 were all obtained from the ATCC. Cells were cultured in DME containing 10% FCS except CHO cells which were maintained in α-MEM/10% FCS at 37 °C in 5% CO_2 until they were near confluence. Washed monolayers of cells (3 X 10^7) were lysed on the dishes and poly A^+ RNA extracted using the Qiagen Oligotex Direct mRNA kit. Samples containing 2 μg of poly A^+ RNA from each cell line were fractionated under denaturing conditions (glyoxal-treated), transferred to a solid nylon membrane support by capillary action, and transcripts visualized by hybridization with random-primed labeled (^{32}P) coding sequence probes derived from either Hu-Asp1 or Hu-Asp2. Radioactive signals were detected by exposure to X-ray film and by image analysis with a PhosphorImager.

The Hu-Asp1 cDNA probe visualized a similar constellation of transcripts (2.6 kb and 3.5 kb) that were previously detected is human tissues. The relative abundance determined by quantification of the radioactive signal was Cos-7 > HEK 292 = HELA > IMR32.

The Hu-Asp2 cDNA probe also visualized a similar constellation of transcripts compared to tissue (3.0 kb, 4.4 kb, and 8.0 kb) with the following relative abundance; HEK 293 > Cos 7 > IMR32 > HELA.

Example 6: Modification of APP to increase Aβ processing for in vitro screening

Human cell lines that process Aβ peptide from APP provide a means to screen in cellular assays for inhibitors of β- and γ-secretase. Production and release of Aβ peptide into the culture supernatant is monitored by an enzyme-linked immunosorbent assay (EIA). Although expression of APP is widespread and both neural and non-neuronal cell lines

process and release Aβ peptide, levels of endogenous APP processing are low and difficult to detect by EIA. Aβ processing can be increased by expressing in transformed cell lines mutations of APP that enhance Aβ processing. We made the serendipitous observation that addition of two lysine residues to the carboxyl terminus of APP695 increases Aβ processing still further. This allowed us to create a transformed cell line that releases Aβ peptide into the culture medium at the remarkable level of 20,000 pg/ml.

Materials And Methods

Materials:

Human embryonic kidney cell line 293 (HEK293 cells) were obtained internally. The vector pIRES-EGFP was purchased from Clontech. Oligonucleotides for mutation using the polymerase chain reaction (PCR) were purchased from Genosys. A plasmid containing human APP695 (SEQ ID No. 9 [nucleotide] and SEQ ID No. 10 [amino acid]) was obtained from Northwestern University Medical School. This was subcloned into pSK (Stratagene) at the *Not*1 site creating the plasmid pAPP695.

Mutagenesis protocol:

The Swedish mutation (K670N, M671L) was introduced into pAPP695 using the Stratagene Quick Change Mutagenesis Kit to create the plasmid pAPP695NL (SEQ ID No. 11 [nucleotide] and SEQ ID No. 12 [amino acid]). To introduce a di-lysine motif at the C-terminus of APP695, the forward primer #276 5' GACTGACCACTCGACCAGGTTC (SEQ ID No. 47) was used with the "patch" primer #274 5' CGAATTAAATTCCAGCACACTGGCTACTTCTTGTTCTGCATCTCAAAGAAC (SEQ ID No. 48) and the flanking primer #275 CGAATTAAATTCCAGCACACTGGCTA (SEQ ID No. 49) to modify the 3' end of the APP695 cDNA (SEQ ID No. 15 [nucleotide] and SEQ ID No. 16 [amino acid]). This also added a BstX1 restriction site that will be compatible with the BstX1 site in the multiple cloning site of pIRES-EGFP. PCR amplification was performed with a Clontech HF Advantage cDNA PCR kit using the polymerase mix and buffers supplied by the manufacturer. For "patch" PCR, the patch primer was used at 1/20th the molar concentration of the flanking primers. PCR amplification products were purified using a QIAquick PCR purification kit (Qiagen). After digestion with restriction enzymes, products were separated on 0.8% agarose gels and then excised DNA fragments were purified using a QIAquick gel extraction kit (Qiagen).

To reassemble a modified APP695-Sw cDNA, the 5' Not1-Bgl2 fragment of the APP695-Sw cDNA and the 3' Bgl2-BstX1 APP695 cDNA fragment obtained by PCR were

ligated into pIRES-EGFP plasmid DNA opened at the Not1 and BstX1 sites. Ligations were performed for 5 minutes at room temperature using a Rapid DNA Ligation kit (Boehringer Mannheim) and transformed into Library Efficiency DH5a Competent Cells (GibcoBRL Life Technologies). Bacterial colonies were screened for inserts by PCR amplification using primers #276 and #275. Plasmid DNA was purified for mammalian cell transfection using a QIAprep Spin Miniprep kit (Qiagen). The construct obtained was designated pMG125.3 (APPSW-KK, SEQ ID No. 17 [nucleotide] and SEQ ID No. 18 [amino acid]).

Mammalian Cell Transfection:

HEK293 cells for transfection were grown to 80% confluence in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal bovine serum. Cotransfections were performed using LipofectAmine (Gibco-BRL) with 3 μg pMG125.3 DNA and 9 μg pcDNA3.1 DNA per 10 x 10^6 cells. Three days posttransfection, cells were passaged into medium containing G418 at a concentration of 400 μg/ml. After three days growth in selective medium, cells were sorted by their fluorescence.

Clonal Selection of 125.3 cells by FACS:

Cell samples were analyzed on an EPICS Elite ESP flow cytometer (Coulter, Hialeah, FL) equipped with a 488 nm excitation line supplied by an air-cooled argon laser. EGFP emission was measured through a 525 nm band-pass filter and fluorescence intensity was displayed on a 4-decade log scale after gating on viable cells as determined by forward and right angle light scatter. Single green cells were separated into each well of one 96 well plate containing growth medium without G418. After a four day recovery period, G418 was added to the medium to a final concentration of 400 μg/ml. After selection, 32% of the wells contained expanding clones. Wells with clones were expanded from the 96 well plate to a 24 well plate and then a 6 well plate with the fastest growing colonies chosen for expansion at each passage. The final cell line selected was the fastest growing of the final six passaged. This clone, designated 125.3, has been maintained in G418 at 400 ug/ml with passage every four days into fresh medium. No loss of Aβ production of EGFP fluorescence has been seen over 23 passages.

Aβ EIA Analysis (Double Antibody Sandwich ELISA for hAβ 1-40/42):

Cell culture supernatants harvested 48 hr after transfection were analyzed in a standard Aβ EIA as follows. Human Aβ 1-40 or 1-42 was measured using monoclonal antibody (mAb) 6E10 (Senetek, St. Louis, MO) and biotinylated rabbit antiserum 162 or

164 (New York State Institute for Basic Research, Staten Island, NY) in a double antibody sandwich ELISA. The capture antibody 6E10 is specific to an epitope present on the N-terminal amino acid residues 1-16 of hAβ. The conjugated detecting antibodies 162 and 164 are specific for hAβ 1-40 and 1-42, respectively. Briefly, a Nunc Maxisorp 96 well immunoplate was coated with 100 μl/well of mAb 6E10 (5μg/ml) diluted in 0.1M carbonate-bicarbonate buffer, pH 9.6 and incubated at 4°C overnight. After washing the plate 3x with 0.01M DPBS (Modified Dulbecco's Phosphate Buffered Saline (0.008M sodium phosphate, 0.002M potassium phosphate, 0.14M sodium chloride, 0.01 M potassium chloride, pH 7.4) from Pierce, Rockford, Il) containing 0.05% of Tween-20 (DPBST), the plate was blocked for 60 min with 200μl of 10% normal sheep serum (Sigma) in 0.01M DPBS to avoid non-specific binding. Human Aβ 1-40 or 1-42 standards 100μl/well (Bachem, Torrance, CA) diluted, from a 1mg/ml stock solution in DMSO, in culture medium was added after washing the plate, as well as 100μl/well of sample, e.g.conditioned medium of transfected cells. The plate was incubated for 2 hours at room temperature and 4°C overnight. The next day, after washing the plate, 100μl/well biotinylated rabbit antiserum 162 1:400 or 164 1:50 diluted in DPBST + 0.5% BSA was added and incubated at room temperature for 1hr 15 min. Following washes, 100μl/well neutravidin-horseradish peroxidase (Pierce, Rockford, Il) diluted 1:10,000 in DPBST was applied and incubated for 1 hr at room temperature. After the last washes 100μl/well of o-phenylnediamine dihydrochloride (Sigma Chemicals, St. Louis, MO) in 50mM citric acid/100mM sodium phosphate buffer (Sigma Chemicals, St. Louis, MO), pH 5.0, was added as substrate and the color development was monitored at 450nm in a kinetic microplate reader for 20 min. using Soft max Pro software. All standards and samples were run in triplicates. The samples with absorbance values falling within the standard curve were extrapolated from the standard curves using Soft max Pro software and expressed in pg/ml culture medium.

Results:

Addition of two lysine residues to the carboxyl terminus of APP695 greatly increases Aβ processing in HEK293 cells as shown by transient expression (Table 1). Addition of the di-lysine motif to APP695 increases Aβ processing to that seen with the APP695 containing the Swedish mutation. Combining the di-lysine motif with the Swedish mutation further increases processing by an additional 2.8 fold.

Cotransformation of HEK293 cells with pMG125.3 and pcDNA3.1 allowed dual selection of transformed cells for G418 resistance and high level expression of EGFP. After clonal selection by FACS, the cell line obtained, produces a remarkable 20,000 pg Aβ peptide per ml of culture medium after growth for 36 hr in 24 well plates. Production of Aβ peptide under various growth conditions is summarized in Table 2.

TABLE 1. Release of Aβ peptide into the culture medium 48 hr after transient transfection of HEK293 cells with the indicated vectors containing wildtype or modified APP. Values tabulated are mean + SD and P-value for pairwise comparison using Student's t-test assuming unequal variances.

APP Construct	Aβ 1-40 peptide (pg/ml)	Fold Increase	P-value
pIRES-EGFP vector	147 + 28	1.0	
wt APP695 (142.3)	194 + 15	1.3	0.051
wt APP695-KK (124.1)	424 + 34	2.8	3 x 10-5
APP695-Sw (143.3)	457 + 65	3.1	2 x 10-3
APP695-SwKK (125.3)	1308 + 98	8.9	3 x 10-4

TABLE 2. Release of Aβ peptide from HEK125.3 cells under various growth conditions.

Type of Culture Plate	Volume of Medium	Duration of Culture	Ab 1-40 (pg/ml)	Ab 1-42 (pg/ml)
24 well plate	400 ul	36 hr	28,036	1,439

Example 7: Antisense oligomer inhibition of Abeta processing in HEK125.3 cells

The sequences of Hu-Asp1 and Hu-Asp2 were provided to Sequitur, Inc (Natick, MA) for selection of targeted sequences and design of 2nd generation chimeric antisense oligomers using proprietary technology (Sequitur Ver. D Pat pending #3002). Antisense oligomers Lot# S644, S645, S646 and S647 were targeted against Asp1. Antisense oligomers Lot# S648, S649, S650 and S651 were targeted against Asp2. Control antisense oligomers Lot# S652, S653, S655, and S674 were targeted against an irrelevant gene and antisense oligomers Lot #S656, S657, S658, and S659 were targeted against a second irrelevant gene.

For transfection with the antisense oligomers, HEK125.3 cells were grown to about 50% confluence in 6 well plates in Minimal Essential Medium (MEM) supplemented with 10% fetal calf serum. A stock solution of oligofectin G (Sequitur Inc., Natick, MA) at 2 mg/ml was diluted to 50 μg/ml in serum free MEM. Separately, the antisense oligomer stock solution at 100 μM was diluted to 800 nM in Opti-MEM (GIBCO-BRL, Grand

Island, NY). The diluted stocks of oligofectin G and antisense oligomer were then mixed at a ratio of 1:1 and incubated at room temperature. After 15 min incubation, the reagent was diluted 10 fold into MEM containing 10% fetal calf serum and 2 ml was added to each well of the 6 well plate after first removing the old medium. After transfection, cells were grown in the continual presence of the oligofectin G/antisense oligomer. To monitor Ab peptide release, 400 µl of conditioned medium was removed periodically from the culture well and replaced with fresh medium beginning 24 hr after transfection. Data reported are from culture supernatants harvested 48 hr after transfection.

Results:

The 16 different antisense oligomers obtained from Sequitur Inc were transfected separately into HEK125.3 cells to determine their affect on Aβ peptide processing. Only antisense oligomers targeted against Asp1 & Asp2 reduced Abeta processing by HEK125.3 cells with those targeted against Asp2 having a greater inhibitory effect. Both Aβ (1-40) and Aβ (1-42) were inhibited by the same degree. In Table 3, percent inhibition is calculated with respect to untransfected cells. Antisense oligomer reagents giving greater than 50% inhibition are marked with an asterisk. Of the reagents tested, 3 of 4 antisense oligomers targeted against ASP1 gave an average 52% inhibition of Aβ 1-40 processing and 47% inhibition of Aβ 1-42 processing. For ASP2, 4 of 4 antisense oligomers gave greater than 50% inhibition with an average inhibition of 62% for Aβ 1-40 processing and 60% for Aβ 1-42 processing.

Table 3. Inhibition of Aβ peptide release from HEK125.3 cells treated with antisense oligomers.

Gene Targeted	Antisense Oligomer	Abeta (1-40)	Abeta (1-42)
Asp1-1	S 644	62%*	56%*
Asp1-2	S 645	41%*	38%*
Asp1-3	S646	52%*	46%*
Asp1-4	S647	6%	25%
Asp2-1	S648	71%*	67%*
Asp2-2	S649	83%*	76%*
Asp2-3	S650	46%*	50%*
Asp2-4	S651	47%*	46%*
Con1-1	S652	13%	18%
Con1-2	S653	35%	30%
Con1-3	S655	9%	18%
Con1-4	S674	29%	18%
Con2-1	S656	12%	18%
Con2-2	S657	16%	19%
Con2-3	S658	8%	35%

Con2-4	S659	3%	18%

Example 8. Demonstration of Hu-Asp2 β- Secretase Activity in Cultured Cells

Several mutations in APP associated with early onset Alzheimer's disease have been shown to alter Aβ peptide processing. These flank the N- and C-terminal cleavage sites that release A□ from APP. These cleavage sites are referred to as the β-secretase and γ-secretase cleavage sites, respectively. Cleavage of APP at the β-secretase site creates a C-terminal fragment of APP containing 99 amino acids of 11,145 daltons molecular weight. The Swedish KM→NL mutation immediately upstream of the β-secretase cleavage site causes a general increase in production of both the 1-40 and 1-42 amino acid forms of A□ peptide. The London VF mutation (V717→F in the APP770 isoform) has little effect on total A□ peptide production, but appears to preferentially increase the percentage of the longer 1-42 amino acid form of A□ peptide by affecting the choice of γ-secretase cleavage site used during APP processing. Thus, we sought to determine if these mutations altered the amount and type of A□ peptide produced by cultured cells cotransfected with a construct directing expression of Hu-Asp2.

Two experiments were performed which demonstrate Hu-Asp2 β-secretase activity in cultured cells. In the first experiment, treatment of HEK125.3 cells with antisense oligomers directed against Hu-Asp2 transcripts as described in Example 7 was found to decrease the amount of the C-terminal fragment of APP created by β-secretase cleavage (CTF99) (Figure 9). This shows that Hu-Asp2 acts directly or indirectly to facilitate β-secretase cleavage. In the second experiment, increased expression of Hu-Asp2 in transfected mouse Neuro2A cells is shown to increase accumulation of the CTF99 β-secretase cleavage fragment (Figure 10). This increase is seen most easily when a mutant APP-KK clone containing a C-terminal di-lysine motif is used for transfection. A further increase is seen when Hu-Asp2 is cotransfected with APP-Sw-KK containing the Swedish mutation KM →NL. The Swedish mutation is known to increase cleavage of APP by the β-secretase.

A second set of experiments demonstrate Hu-Asp2 facilitates γ-secretase activity in cotransfection experiments with human embryonic kidney HEK293 cells. Cotransfection of Hu-Asp2 with an APP-KK clone greatly increases production and release of soluble Aβ1-40 and Aβ1-42 peptides from HEK293 cells. There is a proportionately greater increase in the release of Aβ1-42. A further increase in production of Aβ1-42 is seen when Hu-Asp2 is cotransfected with APP-VF (SEQ ID No. 13 [nucleotide] and SEQ ID No. 14 [amino acid]) or APP-VF-KK SEQ ID No. 19 [nucleotide] and SEQ ID No. 20 [amino acid]) clones containing the London mutation V717→F. The V717→F mutation is known to alter cleavage specificity of the APP γ-secretase such that the preference for cleavage at the Aβ42 site is increased. Thus, Asp2 acts directly or indirectly to facilitate γ-secretase processing of APP at the β42 cleavage site.

Materials

Antibodies 6E10 and 4G8 were purchased from Senetek (St. Louis, MO). Antibody 369 was obtained from the laboratory of Paul Greengard at the Rockefeller University. Antibody C8 was obtained from the laboratory of Dennis Selkoe at the Harvard Medical School and Brigham and Women's Hospital.

APP Constructs used

The APP constructs used for transfection experiments comprised the following

APP	wild-type APP695 (SEQ ID No. 9 and No. 10)
APP-Sw	APP695 containing the Swedish KM→NL mutation (SEQ ID No. 11 and No. 12),
APP-VF	APP695 containing the London V→F mutation (SEQ ID No. 13 and No. 14)
APP-KK	APP695 containing a C-terminal KK motif (SEQ ID No. 15 and No. 16),
APP-Sw-KK	APP695-Sw containing a C-terminal KK motif (SEQ ID No. 17 and No. 18),
APP-VF-KK	APP695-VF containing a C-terminal KK motif (SEQ ID No. 19 and No. 20).

These were inserted into the vector pIRES-EGFP (Clontech, Palo Alto CA) between the *Not*1 and *BstX*1 sites using appropriate linker sequences introduced by PCR.

Transfection of antisense oligomers or plasmid DNA constructs in HEK293 cells, HEK125.3 cells and Neuro-2A cells,

Human embryonic kidney HEK293 cells and mouse Neuro-2a cells were transfected with expression constructs using the Lipofectamine Plus reagent from Gibco/BRL. Cells were seeded in 24 well tissue culture plates to a density of 70-80% confluence. Four wells per plate were transfected with 2 μg DNA (3:1, APP:cotransfectant), 8μl Plus reagent, and 4μl Lipofectamine in OptiMEM. OptiMEM was added to a total volume of 1 ml, distributed 200 μl per well and incubated 3 hours. Care was taken to hold constant the ratios of the two plasmids used for cotransfection as well as the total amount of DNA used in the transfection. The transfection media was replaced with DMEM, 10%FBS, NaPyruvate, with antibiotic/antimycotic and the cells were incubated under normal conditions (37°, 5% CO_2) for 48 hours. The conditioned media were removed to polypropylene tubes and stored at -80°C until assayed for the content of Aβ1-40 and Aβ1-42 by EIA as described in the preceding examples. Transfection of antisense oligomers into HEK125.3 cells was as described in Example 7.

Preparation of cell extracts, Western blot protocol

Cells were harvested after being transfected with plasmid DNA for about 60 hours. First, cells were transferred to 15-ml conical tube from the plate and centrifuged at 1,500 rpm for 5 min to remove the medium. The cell pellets were washed with PBS for one time. We then lysed the cells with lysis buffer (10 mM HEPES, pH 7.9, 150 mM NaCl, 10% glycerol, 1 mM EGTA, 1 mM EDTA, 0.1 mM sodium vanadate and 1% NP-40). The lysed cell mixtures were centrifuged at 5000 rpm and the supernatant was stored at -20°C as the cell extracts. Equal amounts of extracts from HEK125.3 cells transfected with the Asp2 antisense oligomers and controls were precipitated with antibody 369 that recognizes the C-terminus of APP and then CTF99 was detected in the immunoprecipitate with antibody 6E10. The experiment was repeated using C8, a second precipitating antibody that also recognizes the C-terminus of APP. For Western blot of extracts from mouse Neuro-2a cells cotransfected with Hu-Asp2 and APP-KK, APP-Sw-KK, APP-VF-KK or APP-VF, equal amounts of cell extracts were electrophoresed through 4-10% or 10-20% Tricine gradient gels (NOVEX, San Diego, CA). Full length APP and the CTF99 β-secretase product were detected with antibody 6E10.

Results

Transfection of HEK125.3 cells with Asp2-1 or Asp2-2 antisense oligomers reduces production of the CTF β-secretase product in comparison to cells similarly transfected with control oligomers having the reverse sequence (Asp2-1 reverse & Asp2-2 reverse) In cotransfection experiments, cotransfection of Hu-Asp2 into mouse Neuro-2a cells with the APP-KK construct increased the formation of CTF99. This was further increased if Hu-Asp2 was coexpressed with APP-Sw-KK, a mutant form of APP containing the Swedish KM→NL mutation that increases β-secretase processing.

Cotransfection of Hu-Asp2 with APP has little effect on Aβ40 production but increases Aβ42 production above background (Table 4). Addition of the di-lysine motif to the C-terminus of APP increases Aβ peptide processing about two fold, although Aβ40 and Aβ42 production remain quite low (352 pg/ml and 21 pg/ml, respectively). Cotransfection of Asp2 with APP-KK further increases both Aβ40 and Aβ42 production. The stimulation of Aβ40 production by Hu-Asp2 is more that 3 fold, while production of Aβ42 increases by more than 10 fold. Thus, cotransfection of Hu-Asp2 and APP-KK constructs preferentially increases Aβ42 production.

The APP V717→F mutation has been shown to increase γ-secretase processing at the Aβ42 cleavage site. Cotransfection of Hu-Asp2 with the APP-VF or APP-VF-KK constructs increased Aβ42 production (a two fold increase with APP-VF and a four-fold increase with APP-VF-KK, Table 4), but had mixed effects on Aβ40 production (a slight decrease with APP-VF, and a two fold increase with APP-VF-KK in comparison to the pcDNA cotransfection control. Thus, the effect of Asp2 on Aβ42 production was proportionately greater leading to an increase in the ratio of Aβ42/total Aβ. Indeed, the ratio of Aβ42/total Aβ reaches a very high value of 42% in HEK293 cells cotransfected with Hu-Asp2 and APP-VF-KK.

Western blot showing reduction of CTF99 production by HEK125.3 cells transfected with antisense oligomers targeting the Hu-Asp2 mRNA. (right) Western blot showing increase in CTF99 production in mouse Neuro-2a cells cotransfected with Hu-Asp2 and APP-KK. A further increase in CTF99 production is seen in cells cotransfected with Hu-Asp2 and APP-Sw-KK.

Table 4. Results of cotransfecting Hu-Asp2 or pcDNA plasmid DNA with various APP constructs containing the V717→F mutation that modifies γ-secretase processing. Cotransfection with Asp2 consistently increases the ratio of Aβ42/total Aβ. Values tabulated are Aβ peptide pg/ml.

		pcDNA Cotransfection		Asp2 Cotransfection		
	Aβ40	Aβ42	Aβ42/Total	Aβ40	Aβ42	Aβ42/Total
APP	192±18	<4	<2%	188±40	8±10	3.9%
APP-VF	118±15	15±19	11.5%	85±7	24±12	22.4%
APP-KK	352±24	21±6	5.5%	1062±101	226±49	17.5%
APP-VF-KK	230±31	88±24	27.7%	491±35	355±36	42%

Example 9. Bacterial expression of human Asp2L

Expression of recombinant Hu_Asp2L in E. coli.

Hu-Asp2L can be expressed in E. coli after addition of N-terminal sequences such as a T7 tag (SEQ ID No. 21 and No. 22) or a T7 tag followed by a caspase 8 leader sequence (SEQ ID No. 23 and No. 24). Alternatively, reduction of the GC content of the 5' sequence by site directed mutagenesis can be used to increase the yield of Hu-Asp2 (SEQ ID No. 25 and No. 26). In addition, Asp2 can be engineered with a proteolytic cleavage site (SEQ ID No. 27 and No. 28). To produce a soluble protein after expression and refolding, deletion of the transmembrane domain and cytoplasmic tail, or deletion of the membrane proximal region, transmembrane domain, and cytoplasmic tail is preferred.

Methods

PCR with primers containing appropriate linker sequences was used to assemble fusions of Asp2 coding sequence with N-terminal sequence modifications including a T7 tag (SEQ ID Nos. 21 and 22) or a T7-caspase 8 leader (SEQ ID Nos. 23 and 24). These constructs were cloned into the expression vector pet23a(+) [Novagen] in which a T7 promoter directs expression of a T7 tag preceding a sequence of multiple cloning sites. To clone Hu-Asp2 sequences behind the T7 leader of pet23a+, the following oligonucleotides were used for amplification of the selected Hu-Asp2 sequence:
#553=GTGGATCCACCCAGCACGGCATCCGGCTG (SEQ ID No. 35),
#554=GAAAGCTTTCATGACTCATCTGTCTGTGGAATGTTG (SEQ ID No. 36) which placed BamHI and HindIII sites flanking the 5' and 3' ends of the insert, respectively. The Asp2 sequence was amplified from the full length Asp2(b) cDNA cloned into pcDNA3.1 using the Advantage-GC cDNA PCR [Clontech] following the manufacturer's supplied protocol using annealing & extension at 68°C in a two-step PCR cycle for 25 cycles. The insert and vector were cut with BamHI and HindIII, purified by electrophoresis through an agarose gel, then ligated using the Rapid DNA Ligation kit [Boerhinger Mannheim]. The ligation reaction was used to transform the E. coli strain JM109 (Promega) and colonies were picked for the purification of plasmid (Qiagen,Qiaprep minispin) and DNA sequence analysis . For inducible expression using induction with isopropyl b-D-thiogalactopyranoside (IPTG), the expression vector was transferred into E. coli strain BL21 (Statagene). Bacterial cultures were grown in LB broth in the presence of ampicillin at 100 ug/ml, and induced in log phase growth at an OD600 of 0.6-1.0 with 1 mM IPTG for 4 hour at 37°C. The cell pellet was harvested by centrifugation.

To clone Hu-Asp2 sequences behind the T7 tag and caspase leader (SEQ ID Nos. 23 and 24), the construct created above containing the T7-Hu-Asp2 sequence (SEQ ID Nos. 21 and 22) was opened at the BamH1 site, and then the phosphorylated caspase 8 leader oligonucleotides #559=GATCGATGACTATCTCTGACTCTCCGCGTGAACAGGACG (SEQ ID No. 37), #560=GATCCGTCCTGTTCACGCGGAGAGTCAGAGATAGTCATC (SEQ ID No. 38) were annealed and ligated to the vector DNA. The 5' overhang for each set of oligonucleotides was designed such that it allowed ligation into the BamHI site but not subsequent digestion with BamHI. The ligation reaction was transformed into JM109 as above for analysis of protein expression after transfer to E. coli strain BL21.

In order to reduce the GC content of the 5' terminus of asp2, a pair of antiparallel oligos were designed to change degenerate codon bases in 15 amino acid positions from G/C to A/T (SEQ ID Nos. 25 and 26). The new nucleotide sequence at the 5' end of asp2 did not change the encoded amino acid and was chosen to optimize E. Coli expression. The sequence of the sense linker is 5' CGGCATCCGGCTGCCCCTGCGTAGCGGTCTGGGTGGTGCTCCACTGGGTCTGCG TCTGCCCCGGGAGACCGACGAA G 3' (SEQ ID No. 39). The sequence of the antisense linker is : 5' CTTCGTCGGTCTCCCGGGGCAGACGCAGACCCAGTGGAGCACCACCCAGACCG CTACGCAGGGGCAGCCGGATGCCG 3' (SEQ ID No. 40). After annealing the phosphorylated linkers together in 0.1 M NaCl-10 mM Tris, pH 7.4 they were ligated into unique Cla I and Sma I sites in Hu-Asp2 in the vector pTAC. For inducible expression using induction with isopropyl b-D-thiogalactopyranoside (IPTG), bacterial cultures were grown in LB broth in the presence of ampicillin at 100 ug/ml, and induced in log phase growth at an OD600 of 0.6-1.0 with 1 mM IPTG for 4 hour at 37°C. The cell pellet was harvested by centrifugation.

To create a vector in which the leader sequences can be removed by limited proteolysis with caspase 8 such that this liberates a Hu-Asp2 polypeptide beginning with the N-terminal sequence GSFV (SEQ ID Nos. 27 and 28), the following procedure was followed. Two phosphorylated oligonucleotides containing the caspase 8 cleavage site IETD, #571=5' GATCGATGACTATCTCTGACTCTCCGCTGGACTCTGGTATCGAAACCGACG (SEQ ID No. 41) and #572= GATCCGTCGGTTTCGATACCAGAGTCCAGCGGAGAGTCAGAGATAGTCATC (SEQ ID No. 42) were annealed and ligated into pET23a+ that had been opened with BamHI. After transformation into JM109, the purified vector DNA was recovered and orientation of the insert was confirmed by DNA sequence analysis. +, the following oligonucleotides were used for amplification of the selected Hu-Asp2 sequence: #573=5'AAGGATCCTTTGTGGAGATGGTGGACAACCTG, (SEQ ID No. 43) #554=GAAAGCTTTCATGACTCATCTGTCTGTGGAATGTTG (SEQ ID No. 44) which placed BamHI and HindIII sites flanking the 5' and 3' ends of the insert, respectively. The Asp2 sequence was amplified from the full length Asp2 cDNA cloned into pcDNA3.1 using the Advantage-GC cDNA PCR [Clontech] following the manufacturer's supplied

protocol using annealing & extension at 68°C in a two-step PCR cycle for 25 cycles. The insert and vector were cut with BamHI and HindIII, purified by electrophoresis through an agarose gel, then ligated using the Rapid DNA Ligation kit [Boerhinger Mannheim]. The ligation reaction was used to transform the E. coli strain JM109 [Promega] and colonies were picked for the purification of plasmid (Qiagen,Qiaprep minispin) and DNA sequence analysis . For inducible expression using induction with isopropyl b-D-thiogalactopyranoside (IPTG), the expression vector was transferred into E. coli strain BL21 (Statagene). Bacterial cultures were grown in LB broth in the presence of ampicillin at 100 ug/ml, and induced in log phase growth at an OD600 of 0.6-1.0 with 1 mM IPTG for 4 hour at 37°C. The cell pellet was harvested by centrifugation.

To assist purification, a 6-His tag can be introduced into any of the above constructs following the T7 leader by opening the construct at the BamHI site and then ligating in the annealed, phosphorylated oligonucleotides containing the six histidine sequence #565=GATCGCATCATCACCATCACCATG (SEQ ID No. 45), #566=GATCCATGGTGATGGTGATGATGC (SEQ ID No. 46). The 5' overhang for each set of oligonucleotides was designed such that it allowed ligation into the BamHI site but not subsequent digestion with BamHI.

Preparation of Bacterial Pellet:

36.34g of bacterial pellet representing 10.8L of growth was dispersed into a total volume of 200ml using a 20mm tissue homogenizer probe at 3000 to 5000 rpm in 2M KCl, 0.1M Tris, 0.05M EDTA, 1mM DTT. The conductivity adjusted to about 193mMhos with water.

After the pellet was dispersed, an additional amount of the KCl solution was added, bringing the total volume to 500 ml. This suspension was homogenized further for about 3 minutes at 5000 rpm using the same probe. The mixture was then passed through a Rannie high-pressure homogenizer at 10,000psi.

In all cases, the pellet material was carried forward, while the soluble fraction was discarded.The resultant solution was centrifuged in a GSA rotor for 1hr. at 12,500 rpm. The pellet was resuspended in the same solution (without the DTT) using the same tissue homogenizer probe at 2,000 rpm. After homogenizing for 5 minutes at 3000 rpm, the volume was adjusted to 500ml with the same solution, and spun for 1hr. at 12,500 rpm. The pellet was then resuspended as before, but this time the final volume was adjusted to

1.5L with the same solution prior to homogenizing for 5 minutes. After centrifuging at the same speed for 30 minutes, this procedure was repeated. The pellet was then resuspended into about 150ml of cold water, pooling the pellets from the six centrifuge tubes used in the GSA rotor. The pellet has homogenized for 5 minutes at 3,000 rpm, volume adjusted to 250ml with cold water, then spun for 30 minutes. Weight of the resultant pellet was 17.75g.

Summary: Lysis of bacterial pellet in KCl solution, followed by centrifugation in a GSA rotor was used to initially prepare the pellet. The same solution was then used an additional three times for resuspension/homogenization. A final water wash/homogenization was then performed to remove excess KCl and EDTA.

Solublization of rHuAsp2L:

A ratio of 9-10ml/gram of pellet was utilized for solubilizing the rHuAsp2L from the pellet previously described. 17.75g of pellet was thawed, and 150ml of 8M guanidine HCl, 5mM βME, 0.1% DEA, was added. 3M Tris was used to titrate the pH to 8.6. The pellet was initially resuspended into the guanidine solution using a 20mm tissue homogenizer probe at 1000 rpm. The mixture was then stirred at 4°C for 1 hour prior to centrifugation at 12,500rpm for 1 hour in GSA rotor. The resultant supernatant was then centrifuged for 30min at 40,000 x g in an SS-34 rotor. The final supernatant was then stored at –20°C, except for 50ml.

Immobilized Nickel Affinity Chromatography of Solubilized rHuAsp2L:

The following solutions were utilized:

A) 6M Guanidine HCl, 0.1M NaP, pH 8.0, 0.01M Tris, 5mM βME, 0.5mM Imidazole
A') 6M Urea, 20mM NaP, pH 6.80, 50mM NaCl
B') 6M Urea, 20mM NaP, pH 6.20, 50mM NaCl, 12mM Imidazole
C') 6M Urea, 20mM NaP, pH 6.80, 50mM NaCl, 300mM Imidazole

Note: Buffers A' and C' were mixed at the appropriate ratios to give intermediate concentrations of Imidazole.

The 50ml of solubilized material was combined with 50ml of buffer A prior to adding to 100-125ml Qiagen Ni-NTA SuperFlow (pre-equilibrated with buffer A) in a 5 x 10cm Bio-Rad econo column. This was shaken gently overnight at 4°C in the cold room.

Chromatography Steps:

1) Drained the resultant flow through.
2) Washed with 50ml buffer A (collecting into flow through fraction)
3) Washed with 250ml buffer A (wash 1)
4) Washed with 250ml buffer A (wash 2)
5) Washed with 250ml buffer A'

6) Washed with 250ml buffer B'
7) Washed with 250ml buffer A'
8) Eluted with 250ml 75mM Imidazole
9) Eluted with 250ml 150mM Imidazole (150-1)
10) Eluted with 250ml 150mM Imidazole (150-2)
11) Eluted with 250ml 300mM Imidazole (300-1)
12) Eluted with 250ml 300mM Imidazole (300-2)
13) Eluted with 250ml 300mM Imidazole (300-3)

Chromatography Results:

The rHuAsp eluted at 75mM Imidazole through 300mM Imidazole. The 75mM fraction, as well as the first 150mM Imidazole (150-1) fraction contained contaminating proteins as visualized on Coomassie Blue stained gels. Therefore, fractions 150-2 and 300-1 will be utilized for refolding experiments since they contained the greatest amount of protein (see Coomassie Blue stained gel).

Refolding Experiments of rHuAsp2L:

Experiment 1:

Forty ml of 150-2 was spiked with 1M DTT, 3M Tris, pH 7.4 and DEA to a final concentration of 6mM, 50mM, and 0.1% respectively. This was diluted suddenly (while stirring) with 200ml of (4°C) cold 20mM NaP, pH 6.8, 150mM NaCl. This dilution gave a final Urea concentration of 1M. This solution remained clear, even if allowed to set open to the air at RT or at 4°C.

After setting open to the air for 4-5 hours at 4°C, this solution was then dialyzed overnight against 20mM NaP, pH 7.4, 150mM NaCl, 20% glycerol. This method effectively removes the urea in the solution without precipitation of the protein.

Experiment 2:

Some of the 150-2 eluate was concentrated 2x on an Amicon Centriprep, 10,000 MWCO, then treated as in Experiment 1. This material also stayed in solution, with no visible precipitation.

Experiment 3:

89ml of the 150-2 eluate was spiked with 1M DTT, 3M Tris, pH 7.4 and DEA to a final concentration of 6mM, 50mM, and 0.1% respectively. This was diluted suddenly (while stirring) with 445ml of (4°C) cold 20mM NaP, pH 6.8, 150mM NaCl. This solution appeared clear, with no apparent precipitation. The solution was removed to RT and stirred for 10 minutes prior to adding MEA to a final concentration of 0.1mM. This was stirred slowly at RT for 1hr. Cystamine and $CuSO_4$ were then added to final concentrations of 1mM and 10μM respectively. The solution was stirred slowly at RT for 10 minutes prior to being moved to the 4°C cold room and shaken slowly overnight, open to the air.

The following day, the solution (still clear, with no apparent precipitation) was centrifuged at 100,000 x g for 1 hour. Supernatants from multiple runs were pooled, and the bulk of the stabilized protein was dialyzed against 20mM NaP, pH 7.4, 150mM NaCl, 20% glycerol. After dialysis, the material was stored at -20°C.

Some (about 10ml) of the protein solution (still in 1M Urea) was saved back for biochemical analyses, and frozen at -20°C for storage.

Example 10. Expression of Hu-Asp2 and Derivatives in Insect Cells

Expression by baculovirus infection—The coding sequence of Hu-Asp2 and several derivatives were engineered for expression in insect cells using the PCR. For the full-length sequence, a 5'-sense oligonucleotide primer that modified the translation initiation site to fit the Kozak consensus sequence was paired with a 3'-antisense primer that contains the natural translation termination codon in the Hu-Asp2 sequence. PCR amplification of the pcDNA3.1(hygro)/Hu-Asp2 template (see Example 12). Two derivatives of Hu-Asp2 that delete the C-terminal transmembrane domain (SEQ ID No. 29 and No. 30) or delete the transmembrane domain and introduce a hexa-histidine tag at the C-terminus (SEQ ID No. 31 and No. 32) were also engineered using the PCR. The same 5'-sense oligonucleotide primer described above was paired with either a 3'-antisense primer that (1) introduced a translation termination codon after codon 453 (SEQ ID No. 3) or (2) incorporated a hexa-histidine tag followed by a translation termination codon in the PCR using pcDNA3.1(hygro)/Hu_Asp-2L as the template. In all cases, the PCR reactions were performed amplified for 15 cycles using *Pwo*I DNA polymerase (Boehringer-Mannheim) as outlined by the supplier. The reaction products were digested to completion with *BamH*I and *Not*I and ligated to *BamH*I and *Not*I digested baculovirus transfer vector pVL1393 (Invitrogen). A portion of the ligations was used to transform competent *E. coli* DH5α cells

followed by antibiotic selection on LB-Amp. Plasmid DNA was prepared by standard alkaline lysis and banding in CsCl to yield the baculovirus transfer vectors pVL1393/Asp2, pVL1393/Asp2ΔTM and pVL1393/Asp2ΔTM$(His)_6$. Creation of recombinant baculoviruses and infection of sf9 insect cells was performed using standard methods.

Expression by transfection—Transient and stable expression of Hu-Asp2ΔTM and Hu-Asp2ΔTM$(His)_6$ in High 5 insect cells was performed using the insect expression vector pIZ/V5-His. The DNA inserts from the expression plasmids vectors pVL1393/Asp2, pVL1393/Asp2ΔTM and pVL1393/Asp2ΔTM$(His)_6$ were excised by double digestion with *BamHI* and *NotI* and subcloned into *BamHI* and *NotI* digested pIZ/V5-His using standard methods. The resulting expression plasmids, referred to as pIZ/Hu-Asp2ΔTM and pIZ/Hu-Asp2ΔTM$(His)_6$, were prepared as described above.

For transfection, High 5 insect cells were cultured in High Five serum free medium supplemented with 10 μg/ml gentamycin at 27 °C in sealed flasks. Transfections were performed using High five cells, High five serum free media supplemented with 10 μg/ml gentamycin, and InsectinPlus liposomes (Invitrogen, Carlsbad, CA) using standard methods.

For large scale transient transfections 1.2 x 10^7 high five cells were plated in a 150 mm tissue culture dish and allowed to attach at room temperature for 15-30 minutes. During the attachment time the DNA/ liposome mixture was prepared by mixing 6 ml of serum free media, 60 μg Asp2ΔTM/pIZ (+/- His) DNA and 120 μl of Insectin Plus and incubating at room temperature for 15 minutes. The plating media was removed from the dish of cells and replaced with the DNA/liposome mixture for 4 hours at room temperature with constant rocking at 2 rpm. An additional 6 ml of media was added to the dish prior to incubation for 4 days at 27 °C in a humid incubator. Four days post transfection the media was harvested, clarified by centrifugation at 500 x g, assayed for Asp2 expression by Western blotting. For stable expression, the cells were treated with 50 μg/ml Zeocin and the surviving pool used to prepared clonal cells by limiting dilution followed by analysis of the expression level as noted above.

Purification of Hu-Asp2ΔTM and Hu-Asp2ΔTM$(His)_6$—Removal of the transmembrane segment from Hu-Asp2 resulted in the secretion of the polypeptide into the culture medium. Following protein production by either baculovirus infection or transfection, the conditioned medium was harvested, clarified by centrifugation, and dialyzed against Tris-HCl (pH 8.0). This material was then purified by sucessive

chromatography by anion exchange (Tris-HCl, pH 8.0) followed by cation exchange chromatography (Acetate buffer at pH 4.5) using NaCl gradients. The elution profile was monitored by (1) Western blot analysis and (2) by activity assay using the peptide substrate described in Example 12. For the Hu-Asp2ΔTM(His)$_6$, the conditioned medium was dialyzed against Tris buffer (pH 8.0) and purified by sequential chromatography on IMAC resin followed by anion exchange chromatography.

Sequence analysis of the purified Hu-Asp2ΔTM(His)$_6$ protein revealed that the signal peptide had been cleaved [TQHGIRLPLR].

Example 11. Expression of Hu-Asp2 in CHO cells

Heterologous expression of Hu_Asp-2L in CHO-K1 cells—The entire coding sequence of Hu-Asp2 was cloned into the mammalian expression vector pcDNA3.1(+)Hygro (Invitrogen, Carlsbad, CA) which contains the CMV immediate early promotor and bGH polyadenylation signal to drive over expression. The expression plasmid, pcDNA3.1(+)Hygro/Hu-Asp2, was prepared by alkaline lysis and banding in CsCl and completely sequenced on both strands to verify the integrity of the coding sequence.

Wild-type Chinese hamster ovary cells (CHO-K1) were obtained from the ATCC. The cells were maintained in monolayer cultures in α-MEM containing 10% FCS at 37°C in 5% CO_2. Two 100 mm dishes of CHO-K1 cells (60% confluent) were transfected with pcDNA3.1(+)/Hygro alone (mock) or pcDNA3.1(+)Hygro/Hu-Asp2 using the cationic liposome DOTAP as recommended by the supplier. The cells were treated with the plasmid DNA/liposome mixtures for 15 hr and then the medium replaced with growth medium containing 500 Units/ml hygromycin B. In the case of pcDNA3.1(+)Hygro/Hu-Asp2 transfected CHO-K1cells, individual hygromycin B-resistant cells were cloned by limiting dilution. Following clonal expansion of the individual cell lines, expression of Hu-Asp2 protein was accessed by Western blot analysis using a polyclonal rabbit antiserum raised

against recombinant Hu-Asp2 prepared by expression in *E. coli*. Near confluent dishes of each cell line were harvested by scraping into PBS and the cells recovered by centrifugation. The cell pellets were resuspended in cold lysis buffer (25 mM Tris-HCl (8.0)/5 mM EDTA) containing protease inhibitors and the cells lysed by sonication. The soluble and membrane fractions were separated by centrifugation (105,000 x g, 60 min) and normalized amounts of protein from each fraction were then separated by SDS-PAGE. Following electrotransfer of the separated polypeptides to PVDF membranes, Hu_Asp-2L protein was detected using rabbit anti-Hu-Asp2 antiserum (1/1000 dilution) and the antibody-antigen complexes were visualized using alkaline phosphatase conjugated goat anti-rabbit antibodies (1/2500). A specific immunoreactive protein with an apparent Mr value of 65 kDa was detected in pcDNA3.1(+)Hygro/Hu-Asp2 transfected cells and not mock-transfected cells. Also, the Hu-Asp2 polypeptide was only detected in the membrane fraction, consistent with the presence of a signal peptide and single transmembrane domain in the predicted sequence. Based on this analysis, clone #5 had the highest expression level of Hu-Asp2 protein and this production cell lines was scaled up to provide material for purification.

Purification of recombinant Hu_Asp-2L from CHO-K1/Hu-Asp2 clone #5—In a typical purification, clone #5 cell pellets derived from 20 150 mm dishes of confluent cells, were used as the starting material. The cell pellets were resuspended in 50 ml cold lysis buffer as described above. The cells were lysed by polytron homogenization (2 x 20 sec) and the lysate centrifuged at 338,000 x g for 20 minutes. The membrane pellet was then resuspended in 20 ml of cold lysis buffer containing 50 mM β-octylglucoside followed by rocking at 4°C for 1hr. The detergent extract was clarified by centrifugation at 338,000 x g for 20 minutes and the supernatant taken for further analysis.

The β-octylglucoside extract was applied to a Mono Q anion exchange column that was previously equilibrated with 25 mM Tris-HCl (pH 8.0)/50 mM β-octylglucoside. Following sample application, the column was eluted with a linear gradient of increasing NaCl concentration (0-1.0 M over 30 minutes) and individual fractions assayed by Western blot analysis and for β-secretase activity (see below). Fractions containing both Hu_Asp-2L immunoreactivity and β-secretase activity were pooled and dialyzed against 25 mM NaOAc (pH 4.5)/50 mM β-octylglucoside. Following dialysis, precipitated material was removed by centrifugation and the soluble material chromatographed on a MonoS cation exchange column that was previously equilibrated in 25 mM NaOAc (pH 4.5)/ 50 mM β-octylglucoside. The column was eluted using a linear gradient of increasing NaCl concentration (0-1.0 M over 30 minutes) and individual fractions assayed by Western blot analysis and for β-secretase activity. Fractions containing both Hu-Asp2 immunoreactivity and β-secretase activity were combined and determined to be >90% pure by SDS-PAGE/Coomassie Blue staining.

Example 12. Assay of Hu-Asp2 β-secretase activity using peptide substrates

β-secretase assay—β-secretase activity was measured by quantifying the hydrolysis of a synthetic peptide containing the APP Swedish mutation by RP-HPLC with UV detection. Each reaction contained 50 mM Na-MES (pH 5.5), 1% β-octylglucoside, peptide substrate (SEVNLDAEFR, 70 μM) and enzyme (1-5 μg protein). Reactions were incubated at 37 °C for various times and the reaction products were resolved by RP-HPLC using a linear gradient from 0-70 B over 30 minutes (A=0.1% TFA in water, B=).1%TFA/10%water/90%AcCN). The elution profile was monitored by absorbance at 214 nm. In preliminary experiments, the two product peaks which eluted before the intact peptide substrate, were confirmed to have the sequence DAEFR and SEVNL using both

Edman sequencing and MADLI-TOF mass spectrometry. Percent hydrolysis of the peptide substrate was calculated by comparing the integrated peak areas for the two product peptides and the starting material derived from the absorbance at 214 nm. The specificity of the protease cleavage reaction was determined by performing the β-secretase assay in the presence of a cocktail of protease inhibitors (8 μM pepstatin A, 10 μM leupeptin, 10 μM E64, and 5 mM EDTA).

An alternative β-secretase assay utilizes internally quenched fluorescent substrates to monitor enzyme activity using fluorescence spectroscopy in a single sample or multiwell format. Each reaction contained 50 mM Na-MES (pH 5.5), peptide substrate MCA-EVKMDAEF[K-DNP] (BioSource International) (50 μM) and purified Hu-Asp-2 enzyme. These components were equilibrated to 37 °C for various times and the reaction initiated by addition of substrate. Excitation was performed at 330 nm and the reaction kinetics were monitored by measuring the fluorescence emission at 390 nm. To detect compounds that modulate Hu-Asp-2 activity, the test compounds were added during the preincubation phase of the reaction and the kinetics of the reaction monitored as described above. Activators are scored as compounds that increase the rate of appearance of fluorescence while inhibitors decrease the rate of appearance of fluorescence.

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples.

Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the invention.

The entire disclosure of all publications cited herein are hereby incorporated by reference.

Claims

What is claimed is:

1. Any isolated or purified nucleic acid polynucleotide that codes for a protease capable of cleaving the beta (β) secretase cleavage site of APP that contains two or more sets of special nucleic acids, where the special nucleic acids are separated by nucleic acids that code for about 100 to 300 amino acid positions, where the amino acids in those positions may be any amino acids, where the first set of special nucleic acids consists of the nucleic acids that code for the peptide DTG, where the first nucleic acid of the first special set of nucleic acids is, the first special nucleic acid, and where the second set of nucleic acids code for either the peptide DSG or DTG, where the last nucleic acid of the second set of nucleic acids is the last special nucleic acid, with the proviso that the nucleic acids disclosed in SEQ ID NO. 1 and SEQ. ID NO. 5 are not included.

2. The nucleic acid polynucleotide of claim 1 where the two sets of nucleic acids are separated by nucleic acids that code for about 125 to 222 amino acid positions, which may be any amino acids.

3. The nucleic acid polynucleotide of claim 2 that code for about 150 to 172 amino acid positions, which may be any amino acids.

4. The nucleic acid polynucleotide of claim that code for about 172 amino acid positions, which may be any amino acids.

5. The nucleic acid polynucleotide of claim 4 where the nucleotides are described in SEQ. ID. NO. 3

6. The nucleic acid polynucleotide of claim 2 where the two sets of nucleic acids are separated by nucleic acids that code for about 150 to 196 amino acid positions.

7. The nucleic acid polynucleotide of claim 6 where the two sets of nucleotides are separated by nucleic acids that code for about 196 amino acids (positions).

8. The nucleic acid polynucleotide of claim 7 where the two sets of nucleic acids are separated by the same nucleic acid sequences that separate the same set of special nucleic acids in SEQ. ID. NO. 5.

9. The nucleic acid polynucleotide of claim 4 where the two sets of nucleic acids are separated by nucleic acids that code for about 150 to 190, amino acid (positions).

10. The nucleic acid polynucleotide of claim 9 where the two sets of nucleotides are separated by nucleic acids that code for about 190 amino acids (positions).

11. The nucleic acid polynucleotide of claim 10 where the two sets of nucleotides are separated by the same nucleic acid sequences that separate the same set of special nucleotides in SEQ. ID. NO. 1.

12. Claims 1-11 where the first nucleic acid of the first special set of amino acids, that is, the first special nucleic acid, is operably linked to any codon where the nuclic acids of that codon codes for any peptide comprising from 1 to 10,000 amino acid (positions).

13. The nucleic acid polynucleotide of claims 1-12 where the first special nucleic acid is operably linked to nucleic acid polymers that code for any peptide selected from the group consisting of: any any reporter proteins or proteins which facilitate purification.

14. The nucleic acid polynucleotide of claims 1-13 where the first special nucleic acid is operably linked to nucleic acid polymers that code for any peptide selected from the group consisting of: immunoglobin-heavy chain, maltose binding protein, glutathion S transfection, Green Fluorescent protein, and ubiquitin.

15. Claims 1-14 where the last nucleic acid of the second set of special amino acids, that is, the last special nucleic acid, is operably linked to nucleic acid polymers that code for any peptide comprising any amino acids from 1 to 10,000 amino acids.

16. Claims 1-15 where the last special nucleic acid is operably linked to any codon linked to nucleic acid polymers that code for any peptide selected from the group consisting of: any reporter proteins or proteins which facilitate purification.

17. The nucleic acid polynucleotide of claims 1-16 where the first special nucleic acid is operably linked to nucleic acid polymers that code for any peptide selected from the group consisting of: immunoglobin-heavy chain, maltose binding protein, glutathion S transfection, Green Fluorescent protein, and ubiquitin.

18. * Any isolated or purified nucleic acid polynucleotide that codes for a protease capable of cleaving the beta secretase cleavage site of APP that contains two or more sets of special nucleic acids, where the special nucleic acids are separated by nucleic acids that code for about 100 to 300 amino acid positions, where the amino acids in those positions may be any amino acids, where the first set of special nucleic acids consists of the nucleic acids that code for DTG, where the first nucleic acid of the first special set of nucleic acids is, the first special nucleic acid, and where the second set of nucleic acids code for either DSG or DTG, where the last nucleic acid of the second set of special nucleic acids is the last special nucleic acid, where the first special nucleic acid is operably linked to nucleic acids that code for any number of amino acids from zero to 81 amino acids and where each of those codons may code for any amino acid.

19. The nucleic acid polynucleotide of claim 18 , where the first special nucleic acid is operably linked to nucleic acids that code for any number of from 64 to 77 amino acids where each codon may code for any amino acid.

20. The nucleic acid polynucleotide of claim 19 , where the first special nucleic acid is operably linked to nucleic acids that code for about 71 amino acids peptide.

21. The nucleic acid polynucleotide of claim 20, where the first special nucleic acid is operably linked to 71 amino acid peptide and where the first of those 71 amino acids is the amino acid T.

22. The nucleic acid polynucleotide of claim 21, where the polynucleotide comprises a sequence that is at least 95% identical to the same corresponding amino acids in SEQ. ID. NO. 3, that is, identical to the sequences in SEQ. ID. NO. 3 including the sequences from both the first and or the second special nucleic acids, toward the N-Terminal, through and including 71 amino acids, see Example 10, beginning from the DTG site and including the nucleotides from that code for 71 amino acids).

23. The nucleic acid polynucleotide of claim 22, where the complete polynucleotide comprises identical to the same corresponding amino acids in SEQ. ID. NO. 3, that is, identical to the sequences in SEQ. ID. NO. 3 including the sequences from both the first and or the second special nucleic acids, toward the N-Terminal, through and including 71 amino acids, see Example 10, beginning from the DTG site and including the nucleotides from that code for 71 amino acids).

24. The nucleic acid polynucleotide of claim 18 , where the first special nucleic acid is operably linked to nucleic acids that code for any number of from about 30 to 54 amino acids where each codon may code for any amino acid.

25. The nucleic acid polynucleotide of claim 20, where the first special nucleic acid is operably linked to 47 codons where the first those 35 or 47 amino acids is the amino acid E or G.

26. The nucleic acid polynucleotide of claim 21, where the polynucleotide comprises a sequence that is at least 95% identical to the same corresponding amino acids in SEQ. ID. NO. 3, that is, identical to that portion of the sequences in SEQ. ID. NO. 3 including the sequences from both the first and or the second special nucleic acids, toward the N-Terminal, through and including 35 or 47 amino acids, see Example 11 for the 47 example, beginning from the DTG site and including the nucleotides from that code for the previous 35 or 47 amino acids before the DTG site).

27. The nucleic acid polynucleotide of claim 22, where the polynucleotide comprises identical to the same corresponding amino acids in SEQ. ID. NO. 3, that is, identical to the sequences in SEQ. ID. NO. 3 including the sequences from both the first and or the second special nucleic acids, toward the N-Terminal, through and including 35 or 47 amino acids, see Example 11 for the 47 example, beginning from the DTG site and including the nucleotides from that code for the previous 35 or 47 amino acids before the DTG site).

28. * Any isolated or purified nucleic acid polynucleotide that codes for a protease capable of cleaving the beta (β) secretase cleavage site of APP that contains two or more sets of special nucleic acids, where the special nucleic acids are separated by nucleic acids that code for about 100 to 300 amino acid positions, where the amino acids in those positions may be any amino acids, where the first set of special nucleic acids consists of the nucleic acids that code for the peptide DTG, where the first nucleic acid of the first special set of amino acids is, the first special nucleic acid, and where the second set of special nucleic acids code for either the peptide DSG or DTG, where the last nucleic acid of the second set of special nucleic acids, the last special nucleic acid, is operably linked to nucleic acids that code for any number of codons from 50 to 170 codons.

29. The nucleic acid polynucleotide of claim 29 where the last special nucleic acid is operably linked to nucleic acids comprising from 100 to 170 codons.

30. The nucleic acid polynucleotide of claim 30 where the last special nucleic acid is operably linked to nucleic acids comprising from 142 to 163 codons.

31. The nucleic acid polynucleotide of claim 31 where the last special nucleic acid is operably linked to nucleic acids comprising about 142 codons.

32. The nucleic acid polynucleotide of claim 32 where the polynucleotide comprises a sequence that is at least 95% identical to SEQ. ID. # (Example 9 or 10).

33. The nucleic acid polynucleotide of claim 33, where the complete polynucleotide comprises SEQ. ID. # (Example 9 or 10).

34. The nucleic acid polynucleotide of claim 31 where the last special nucleic acid is operably linked to nucleic acids comprising about 163 codons.

35. The nucleic acid polynucleotide of claim 35 where the polynucleotide comprises a sequence that is at least 95% identical to SEQ. ID. # (Example 9 or 10).

36. The nucleic acid polynucleotide of claim 36, where the complete polynucleotide comprises SEQ. ID. # (Example 9 or 10).

37. The nucleic acid polynucleotide of claim 31 where the last special nucleic acid is operably linked to nucleic acids comprising about 170 codons.

38. Claims 1-38 where the second set of special nucleid acids code for the peptide DSG, and optionally the first set of nucleic acid polynucleotide is operably linked to a peptide purification tag.

39. Claims 1-39 where the nucleic acid polynucleotide is operably linked to a peptide purification tag which is six histidine.

40. Claims 1-40 where the first set of special nucleic acids are on one polynucleotide and the second set of special nucleic acids are on a second polynucleotide, where both first and second polynucleotides have at lease 50 codons.

41. Claims 1-40 where the first set of special nucleic acids are on one polynucleotide and the second set of special nucleic acids are on a second polynucleotide, where both first and second polynucleotides have at lease 50 codons where both said polynucleotides are in the same solution.

42. A vector which contains a polynucleotide described in claims 1-42.

43. A cell or cell line which contans a polynucleotide described in claims 1-42.

44. Any isolated or purified peptide or protein comprising an amino acid polymer that is a protease capable of cleaving the beta (β) secretase cleavage site of APP that contains two or more sets of special amino acids, where the special amino acids are separated by about 100 to 300 amino acid positions, where each amino acid position can be any amino acid, where the first set of special amino acids consists of the peptide DTG, where the first amino acid of the first special set of amino acids is, the first special amino acid, where the second set of amino acids is selected from the peptide comprising either DSG or DTG, where the last amino acid of the second set of special amino acids is the last special amino acid, with the proviso that the proteases disclosed in SEQ ID NO. 2 and SEQ. ID NO. 6 are not included.

45. The amino acid polypeptide of claim 45 where the two sets of amino acids are separated by about 125 to 222 amino acid positions where in each position it may be any amino acid.

46. The amino acid polypeptide of claim 46 where the two sets of amino acids are separated by about 150 to 172 amino acids.

47. The amino acid polypeptide of claim 47 where the two sets of amino acids are separated by about 172 amino acids.

48. The amino acid polypeptide of claim 48 where the protease is described in SEQ. ID. NO. 4

49. The amino acid polypeptide of claim 46 where the two sets of amino acids are separated by about 150 to 196 amino acids.

50. The amino acid polypeptide of claim 50 where the two sets of amino acids are separated by about 196 amino acids.

51. The amino acid polypeptide of claim 51 where the two sets of amino acids are separated by the same amino acid sequences that separate the same set of special amino acids in SEQ. ID. NO. 6.

52. The amino acid polypeptide of claim 46 where the two sets of amino acids are separated by about 150 to 190. amino acids.

53. The amino acid polypeptide of claim 53 where the two sets of nucleotides are separated by about 190 amino acids.

54. The amino acid polypeptide of claim 54 where the two sets of nucleotides are separated by the same amino acid sequences that separate the same set of special amino acids in SEQ. ID. NO. 2.

55. Claims 45-55 where the first amino acid of the first special set of amino acids, that is, the first special amino acid, is operably linked to any peptide comprising from 1 to 10,000 amino acids.

56. The amino acid polypeptide of claims 45-56 where the first special amino acid is operably linked to any peptide selected from the group consisting of: any any reporter proteins or proteins which facilitate purification.

57. The amino acid polypeptide of claims 45-57 where the first special amino acid is operably linked to any peptide selected from the group consisting of: immunoglobin-heavy chain, maltose binding protein, glutathion S transfection, Green Fluorescent protein, and ubiquitin.

58. Claims 45-58, where the last amino acid of the second set of special amino acids, that is, the last special amino acid, is operably linked to any peptide comprising any amino acids from 1 to 10,000 amino acids.

59. Claims 45-59 where the last special amino acid is operably linked any peptide selected from the group consisting of any reporter proteins or proteins which facilitate purification.

60. The amino acid polypeptide of claims 45-60 where the first special amino acid is operably linked to any peptide selected from the group consisting of: immunoglobin-heavy chain, maltose binding protein, glutathion S transfection, Green Fluorescent protein, and ubiquitin.

61. * Any isolated or purified peptide or protein comprising an amino acid polypeptide that codes for a protease capable of cleaving the beta secretase cleavage site of APP that contains two or more sets of special amino acids, where the special amino acids are separated by about 100 to 300 amino acid positions, where each amino acid in each position can be any amino acid, where the first set of special amino acids consists of the amino acids DTG, where the first amino acid of the first special set of amino acids is, the first special amino acid, D, and where the second set of amino acids is either DSG or DTG, where the last amino acid of the second set of special amino acids is the last special amino acid, G, where the first special amino acid is operably linked to amino acids that code for any number of amino acids from zero to 81 amino acid positions where in each position it may be any amino acid.

62. The amino acid polypeptide of claim 62, where the first special amino acid is operably linked to a peptide from about 30 to 77 amino acids positions where each amino acid position may be any amino acid.

63. The amino acid polypeptide of claim 63, where the first special amino acid is operably linked to a peptide of 35, 47, 71, or 77 amino acids.

64. The amino acid polypeptide of claim 63, where the first special amino acid is operably linked to the same corresponding peptides from SEQ. ID. NO. 3 that are 35, 47, 71, or 77 peptides in length, beginning counting with the amino acids on the first special sequence, DTG, towards the N-terminal of SEQ. ID. NO. 3.

65. The amino acid polypeptide of claim 65, where the polypeptide comprises a sequence that is at least 95% identical to the same corresponding amino acids in SEQ. ID. NO. 4, that is, identical to that portion of the sequences in SEQ.ID. NO. 4, including all the sequences from both the first and or the second special nucleic acids, toward the N- terminal, through and including 71, 47, 35 amino acids before the first special amino acids. (Examples 10 and 11).

66. The amino acid polypeptide of claim 65, where the complete polypeptide comprises the peptide of 71 amino acids, where the first of the amino acid is T and the second is Q.

67. The amino acid polypeptide of claim 62, where the first special amino acid is operably linked to any number of from 40 to 54 amino acids (positions) where each amino acid position may be any amino acid.

68. The amino acid polypeptide of claim 68, where the first special amino acid is operably linked to amino acids that code for a peptide of 47 amino acids.

69. The amino acid polypeptide of claim 69, where the first special amino acid is operably linked to a 47 amino acid peptide where the first those 47 amino acids is the amino acid E.

70. The amino acid polypeptide of claim 70, where the polypeptide comprises a sequence that is at least 95% identical to SEQ. ID. # (Example 10).

71. The amino acid polypeptide of claim 71, where the complete polypeptide comprises SEQ. ID. # (Example 10).

72. * Any isolated or purified amino acid polypeptide that is a protease capable of cleaving the beta (β) secretase cleavage site of APP that contains two or more sets of special amino acids, where the special amino acids are separated by about 100 to 300 amino acid positions, where each amino acid in each position can be any amino

acid, where the first set of special amino acids consists of the amino acids that code for DTG, where the first amino acid of the first special set of amino acids is, the first special amino acid, D, and where the second set of amino acids are either DSG or DTG, where the last amino acid of the second set of special amino acids is the last special amino acid, G, which is operably linked to any number of amino acids from 50 to 170 amino acids, which may be any amino acids.

73. The amino acid polypeptide of claim 73 where the last special amino acid is operably linked to a peptide of about 100 to 170 amino acids.

74. The amino acid polypeptide of claim 74 where the last special amino acid is operably linked to to a peptide of about 142 to 163 amino acids.

75. The amino acid polypeptide of claim 75 where the last special amino acid is operably linked to to a peptide of about about 142 amino acids.

76. The amino acid polypeptide of claim 76 where the polypeptide comprises a sequence that is at least 95% identical to SEQ. ID. # (Example 9 or 10).

77. The amino acid polypeptide of claim 75 where the last special amino acid is operably linked to a peptide of about 163 amino acids.

78. The amino acid polypeptide of claim 79 where the polypeptide comprises a sequence that is at least 95% identical to SEQ. ID. # (Example 9 or 10).

79. The amino acid polypeptide of claim 79, where the complete polypeptide comprises SEQ. ID. # (Example 9 or 10).

80. The amino acid polypeptide of claim 74 where the last special amino acid is operably linked to to a peptide of about 170 amino acids.

81. Claim 46-81 where the second set of special amino acids is comprised of the peptide with the amino acid sequence DSG.

82. Claims 45-82 where the amino acid polypeptide is operably linked to a peptide purification tag.

83. Claims 45-83 where the amino acid polypeptide is operably linked to a peptide purification tag which is six histidine.

84. Claims 45-84 where the first set of special amino acids are on one polypeptide and the second set of special amino acids are on a second polypeptide, where both first and second polypeptide have at lease 50 amino acids, which may be any amino acids.

85. Claims 45-84 where the first set of special amino acids are on one polypeptide and the second set of special amino acids are on a second polypeptide, where both first and second polypeptides have at lease 50 amino acids where both said polypeptides are in the same vessel.

86. A vector which contains a polypeptide described in claims 45-86.

87. A cell or cell line which contans a polynucleotide described in claims 45-87.

88. The process of making any of the polynucleotides, vectors, or cells of claims 1-44

89. The process of making any of the polypeptides, vectors or cells of claims 45-88

90. Any of the polynucleotides, polypeptides, vectors, cells or cell lines described in claims 1-88 made from the processes described in claims 89 and 90.

91. * An isolated nucleic acid molecule comprising a polynucleotide, said polynucleotide encoding a Hu-Asp polypeptide and having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:

(a) a nucleotide sequence encoding a Hu-Asp polypeptide selected from the group consisting of Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b), wherein said Hu-Asp1, Hu-Asp2(a) and Hu-Asp2(b) polypeptides have the complete amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, and SEQ ID No:6, respectively; and

(b) a nucleotide sequence complementary to the nucleotide sequence of (a).

92. The nucleic acid molecule of claim 92, wherein said Hu-Asp polypeptide is Hu-Asp1, and said polynucleotide molecule of 1(a) comprises the nucleotide sequence of SEQ ID NO:1.

93. The nucleic acid molecule of claim 92, wherein said Hu-Asp polypeptide is Hu-Asp2(a), and said polynucleotide molecule of 1(a) comprises the nucleotide sequence of SEQ ID NO:4.

94. The nucleic acid molecule of claim 92, wherein said Hu-Asp polypeptide is Hu-Asp2(b), and said polynucleotide molecule of 1(a) comprises the nucleotide sequence of SEQ ID NO:5.

95. An isolated nucleic acid molecule comprising polynucleotide which hybridizes under stringent conditions to a polynucleotide having the nucleotide sequence in (a) or (b) of claim 92.

96. A vector comprising the nucleic acid molecule of claim 96.

97. The vector of claim 97, wherein said nucleic acid molecule is operably linked to a promoter for the expression of a Hu-Asp polypeptide.

98. The vector of claim 97, wherein said Hu-Asp polypeptide is Hu-Asp1.

99. The vector of claim 97, wherein said Hu-Asp polypeptide is Hu-Asp2(a).

100. The vector of claim 97, wherein said Hu-Asp polypeptide is Hu-Asp2(b).

101. A host cell comprising the vector of claim 98.

102. A method of obtaining a Hu-Asp polypeptide comprising culturing the host cell of claim 102 and isolating said Hu-Asp polypeptide.

103. An isolated Hu-Asp1 polypeptide comprising an amino acid sequence at least 95% identical to a sequence comprising the amino acid sequence of SEQ ID NO:2.

104. An isolated Hu-Asp2(a) polypeptide comprising an amino acid sequence at least 95% identical to a sequence comprising the amino acid sequence of SEQ ID NO:4.

105. An isolated Hu-Asp2(a) polypeptide comprising an amino acid sequence at least 95% identical to a sequence comprising the amino acid sequence of SEQ ID NO:8.

106 An isolated antibody that binds specifically to the Hu-Asp polypeptide of any of claims 104-107.
sequence comprising the amino acid sequence of SEQ ID NO:8.

107 An isolated antibody that binds specifically to the Hu-Asp polypeptide of any of claims 104-107.

108. * A method to identify a cell that can be used to screen for inhibitors of β secretase activity comprising:

a) identifying a cell that expresses a protease capable of cleaving APP at the β secretase site,
comprising:
   i) collect the cells or the supernantent from the cells to be identified
   ii) measure the production of a critical peptide, where the critical peptide is selected from the group consisting of either the APP C-terminal peptide or soluble APP,
   iii) select the cells which produce the critical peptide.

109. The method of claim 108 where the cells are collected and the critical peptide is the APP C-terminal peptide created as a result of the β secretase cleavage.

110. The method of claim 108 where the supernantent is collected and the critical peptide is soluble APP where the soluble APP has a C-terminal created by β secretase cleavage.

111. The method of claim 108 where the cells contain any of the nucleic acids or polypeptides of claims 1-86 and where the cells are shown to cleave the β secretase site of any peptide having the following peptide structure, P2, P1, P1', P2', where P2 is K or N, where P1 is M or L, where P1' is D, where P2' is A.

112. The method of claim 111 where P2 is K and P1 is M.

113 The method of claim 112 where P2 is N and P1 is L.

114 * Any bacterial cell comprising any nucleic acids or peptides in claims 1-86 and 92-107.

115 A bacterial cell of claim 114 where the bacteria is *E coli.*

116 Any eukaryotic cell comprising any nucleic acids or polypeptides in claims 1-86 and 92-107.

117 * Any insect cell comprising any of the nucleic acids or polypeptides in claims 1-86 and 92-107.

118 A insect cell of claim 117 where the insect is sf9, or High 5.

119 A insect cell of claim 100 where the insect cell is High 5.

120 A mammalian cell comprising any of the nucleic acids or polypeptides in claims 1-86 and 92-107.

121 A mammalian cell of claim 120 where the mammalian cell is selected from the group consisting of, human, rodent, lagomorph, and primate.

122 A mammalian cell of claim 121 where the mammalian cell is selected from the group consisting of human cell.

123 A mammalian cell of claim 122 where the human cell is selected from the group comprising HEK293, and IMR-32.

124 A mammalian cell of claim 121 where the cell is a primate cell.

125 A primate cell of claim 124 where the primate cell is a COS-7 cell.

126 A mammalian cell of claim 121 where cell is selected from a rodent cells.

127 A rodent cell of claim 126 selected from, CHO-K1, Neuro-2A, 3T3 cells.

128 A yeast cell of claim 115.

129 An avian cell of claim 115.

130. * Any isoform of APP where the last two carboxy terminus amino acids of that isoform are both lysine residues.

131 The isoform of APP from claim 130 comprising the isoform known as APP695 modified so that its last two having two lysine residues as its last two carboxy terminus amino acids.

132 The isoform of claim 131 comprising SEQ. ID. 16.

133 The isoform variant of claim 1301comprising SEQ. ID. NO. 18, and 20.

134 Any eukaryotic cell line, comprising nucleic acids or polypeptides of claim 130-133.

135 Any cell line of claim 134 that is a mammaliam cell line (HEK293, Neuro2a, are preferred plus any others.)

136 A method for identifying inhibitors of an enzyme that cleaves the beta secretase cleavabe site of APP comprising:

a) culturing cells in a culture medium under conditions in which the enzyme causes processing of APP and release of amyloid beta-peptide into the medium and causes the accumulation of CTF99 fragments of APP in cell lysates,

b) exposing the cultured cells to a test compound; and specifically determining whether the test compound inhibits the function of the enzyme by measuring the amount of amyloid beta-peptide released into the medium and or the amount of CTF99 fragments of APP in cell lysates;

c) identifying test compounds diminishing the amount of soluble amyloid beta peptide present in the culture medium and diminution of CTF99 fragments of APP in cell lysates as Asp2 inhibitors.

137 The method of claim 136 wherein the cultured cells are a human, rodent or insect cell line.

138 The method of claim 137 wherein the human or rodent cell line exhibits β secretase activity in which processing of APP occurs with release of amyloid beta-peptide into the culture medium and accumulation of CTF99 in cell lysates.

139. A method as in claim 138 wherein the human or rodent cell line treated with the antisense oligomers directed against the enzyme that exhibits β secretase activity, reduces release of soluble amyloid beta-peptide into the culture medium and accumulation of CTF99 in cell lysates.

140. A method for the identification of an agent that decreases the activity of a Hu-Asp polypeptide selected from the group consisting of Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b), the method comprising

(a) determining the activity of said Hu-Asp polypeptide in the presence of a test agent and in the absence of a test agent; and

(b) comparing the activity of said Hu-Asp polypeptide determined in the presence of said test agent to the activity of said Hu-Asp polypeptide determined in the absence of said test agent;

whereby a lower level of activity in the presence of said test agent than in the absence of said test agent indicates that said test agent has decreased the activity of said Hu-Asp polypeptide..

141. The nucleic acids, peptides, proteins, vectors, cells and cell lines, and assays described herein.

FIGURE 1 (1)

ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC
M G A L A R A L L L P L L A Q W L L R A

CCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC
A P E L A P A P F T L P L R V A A A T N

CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG
R V V A P T P G P G T P A E R H A D G L

GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG
A L A L E P A L A S P A G A A N F L A M

GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC
V D N L Q G D S G R G Y Y L E M L I G T

CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA
P P Q K L Q I L V D T G S S N F A V A G

ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC
T P H S Y I D T Y F D T E R S S T Y R S

AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA
K G F D V T V K Y T Q G S W T G F V G E

GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT
D L V T I P K G F N T S F L V N I A T I

TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT
F E S E N F F L P G I K W N G I L G L A

TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA
Y A T L A K P S S S L E T F F D S L V T

CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT
Q A N I P N V F S M Q M C G A G L P V A

GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA
G S G T N G G S L V L G G I E P S L Y K

GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA
G D I W Y T P I K E E W Y Y Q I E I L K

TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC
L E I G G Q S L N L D C R E Y N A D K A

ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG
I V D S G T T L L R L P Q K V F D A V V

GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC
E A V A R A S L I P E F S D G F W T G S

CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC
Q L A C W T N S E T P W S Y F P K I S I

TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC
Y L R D E N S S R S F R I T I L P Q L Y

ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA
I Q P M M G A G L N Y E C Y R F G I S P

TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC
S T N A L V I G A T V M E G F Y V I F D

AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA

FIGURE 1 (2)

```
R  A  Q  K  R  V  G  F  A  A  S  P  C  A  E  I  A  G  A  A

GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT
V  S  E  I  S  G  P  F  S  T  E  D  V  A  S  N  C  V  P  A

CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA
Q  S  L  S  E  P  I  L  W  I  V  S  Y  A  L  M  S  V  C  G

GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC
A  I  L  L  V  L  I  V  L  L  L  L  P  F  R  C  Q  R  R  P

CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAATGAATA
R  D  P  E  V  V  N  D  E  S  S  L  V  R  H  R  W  K

GCCAGGCCTGACCTCAAGCAACCATGAACTCAGCTATTAAGAAAATCACATTTCCAGGGC
AGCAGCCGGGATCGATGGTGGCGCTTTCTCCTGTGCCCACCCGTCTTCAATCTCTGTTCT
GCTCCCAGATGCCTTCTAGATTCACTGTCTTTTGATTCTTGATTTTCAAGCTTTCAAATC
CTCCCTACTTCCAAGAAAAATAATTAAAAAAAAAAACTTCATTCTAAACCAAAAAAAAAAA
AAAA
```

FIGURE 2 (1)

```
ATGGCCCAAGCCCTGCCCTGGCTCCTGCTGTGGATGGGCGCGGGAGTGCTGCCTGCCCAC
M  A  Q  A  L  P  W  L  L  L  W  M  G  A  G  V  L  P  A  H

GGCACCCAGCACGGCATCCGGCTGCCCCTGCGCAGCGGCCTGGGGGGCGCCCCCCTGGGG
G  T  Q  H  G  I  R  L  P  L  R  S  G  L  G  G  A  P  L  G

CTGCGGCTGCCCCGGGAGACCGACGAAGAGCCCGAGGAGCCCGGCCGGAGGGGCAGCTTT
L  R  L  P  R  E  T  D  E  E  P  E  E  P  G  R  R  G  S  F

GTGGAGATGGTGGACAACCTGAGGGGCAAGTCGGGGCAGGGCTACTACGTGGAGATGACC
V  E  M  V  D  N  L  R  G  K  S  G  Q  G  Y  Y  V  E  M  T

GTGGGCAGCCCCCCGCAGACGCTCAACATCCTGGTGGATACAGGCAGCAGTAACTTTGCA
V  G  S  P  P  Q  T  L  N  I  L  V  D  T  G  S  S  N  F  A

GTGGGTGCTGCCCCCCACCCCTTCCTGCATCGCTACTACCAGAGGCAGCTGTCCAGCACA
V  G  A  A  P  H  P  F  L  H  R  Y  Y  Q  R  Q  L  S  S  T

TACCGGGACCTCCGGAAGGGTGTGTATGTGCCCTACACCCAGGGCAAGTGGGAAGGGGAG
Y  R  D  L  R  K  G  V  Y  V  P  Y  T  Q  G  K  W  E  G  E

CTGGGCACCGACCTGGTAAGCATCCCCCATGGCCCCAACGTCACTGTGCGTGCCAACATT
L  G  T  D  L  V  S  I  P  H  G  P  N  V  T  V  R  A  N  I

GCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACTGGGAAGGCATCCTG
A  A  I  T  E  S  D  K  F  F  I  N  G  S  N  W  E  G  I  L

GGGCTGGCCTATGCTGAGATTGCCAGGCTTTGTGGTGCTGGCTTCCCCCTCAACCAGTCT
G  L  A  Y  A  E  I  A  R  L  C  G  A  G  F  P  L  N  Q  S

GAAGTGCTGGCCTCTGTCGGAGGGAGCATGATCATTGGAGGTATCGACCACTCGCTGTAC
E  V  L  A  S  V  G  G  S  M  I  I  G  G  I  D  H  S  L  Y

ACAGGCAGTCTCTGGTATACACCCATCCGGCGGGAGTGGTATTATGAGGTGATCATTGTG
T  G  S  L  W  Y  T  P  I  R  R  E  W  Y  Y  E  V  I  I  V

CGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAGTACAACTATGACAAG
R  V  E  I  N  G  Q  D  L  K  M  D  C  K  E  Y  N  Y  D  K

AGCATTGTGGACAGTGGCACCACCAACCTTCGTTTGCCCAAGAAAGTGTTTGAAGCTGCA
S  I  V  D  S  G  T  T  N  L  R  L  P  K  K  V  F  E  A  A

GTCAAATCCATCAAGGCAGCCTCCTCCACGGAGAAGTTCCCTGATGGTTTCTGGCTAGGA
V  K  S  I  K  A  A  S  S  T  E  K  F  P  D  G  F  W  L  G

GAGCAGCTGGTGTGCTGGCAAGCAGGCACCACCCCTTGGAACATTTTCCCAGTCATCTCA
E  Q  L  V  C  W  Q  A  G  T  T  P  W  N  I  F  P  V  I  S

CTCTACCTAATGGGTGAGGTTACCAACCAGTCCTTCCGCATCACCATCCTTCCGCAGCAA
L  Y  L  M  G  E  V  T  N  Q  S  F  R  I  T  I  L  P  Q  Q

TACCTGCGGCCAGTGGAAGATGTGGCCACGTCCCAAGACGACTGTTACAAGTTTGCCATC
```

FIGURE 2 (2)

```
Y  L  R  P  V  E  D  V  A  T  S  Q  D  D  C  Y  K  F  A  I

TCACAGTCATCCACGGGCACTGTTATGGGAGCTGTTATCATGGAGGGCTTCTACGTTGTC
S  Q  S  S  T  G  T  V  M  G  A  V  I  M  E  G  F  Y  V  V

TTTGATCGGGCCCGAAAACGAATTGGCTTTGCTGTCAGCGCTTGCCATGTGCACGATGAG
F  D  R  A  R  K  R  I  G  F  A  V  S  A  C  H  V  H  D  E

TTCAGGACGGCAGCGGTGGAAGGCCCTTTTGTCACCTTGGACATGGAAGACTGTGGCTAC
F  R  T  A  A  V  E  G  P  F  V  T  L  D  M  E  D  C  G  Y

AACATTCCACAGACAGATGAGTCAACCCTCATGACCATAGCCTATGTCATGGCTGCCATC
N  I  P  Q  T  D  E  S  T  L  M  T  I  A  Y  V  M  A  A  I

TGCGCCCTCTTCATGCTGCCACTCTGCCTCATGGTGTGTCAGTGGCGCTGCCTCCGCTGC
C  A  L  F  M  L  P  L  C  L  M  V  C  Q  W  R  C  L  R  C

CTGCGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTGAAGTGAGGAGGCCCA
L  R  Q  Q  H  D  D  F  A  D  D  I  S  L  L  K

TGGGCAGAAGATAGAGATTCCCCTGGACCACACCTCCGTGGTTCACTTTGGTCACAAGTA
GGAGACACAGATGGCACCTGTGGCCAGAGCACCTCAGGACCCTCCCCACCCACCAAATGC
CTCTGCCTTGATGGAGAAGGAAAAGGCTGGCAAGGTGGGTTCCAGGGACTGTACCTGTAG
GAAACAGAAAAGAGAAGAAAGAAGCACTCTGCTGGCGGGAATACTCTTGGTCACCTCAAA
TTTAAGTCGGGAAATTCTGCTGCTTGAAACTTCAGCCCTGAACCTTTGTCCACCATTCCT
TTAAATTCTCCAACCCAAAGTATTCTTCTTTTCTTAGTTTCAGAAGTACTGGCATCACAC
GCAGGTTACCTTGGCGTGTGTCCCTGTGGTACCCTGGCAGAGAAGAGACCAAGCTTGTTT
CCCTGCTGGCCAAAGTCAGTAGGAGAGGATGCACAGTTTGCTATTTGCTTTAGAGACAGG
GACTGTATAAACAAGCCTAACATTGGTGCAAAGATTGCCTCTTGAAAAAAAAAAAAA
```

FIGURE 3 (1)

```
ATGGCCCAAGCCCTGCCCTGGCTCCTGCTGTGGATGGGCGCGGGAGTGCTGCCTGCCCAC
M  A  Q  A  L  P  W  L  L  L  W  M  G  A  G  V  L  P  A  H

GGCACCCAGCACGGCATCCGGCTGCCCCTGCGCAGCGGCCTGGGGGGCGCCCCCCTGGGG
G  T  Q  H  G  I  R  L  P  L  R  S  G  L  G  G  A  P  L  G

CTGCGGCTGCCCCGGGAGACCGACGAAGAGCCCGAGGAGCCCGGCCGGAGGGGCAGCTTT
L  R  L  P  R  E  T  D  E  E  P  E  E  P  G  R  R  G  S  F

GTGGAGATGGTGGACAACCTGAGGGGCAAGTCGGGGCAGGGCTACTACGTGGAGATGACC
V  E  M  V  D  N  L  R  G  K  S  G  Q  G  Y  Y  V  E  M  T

GTGGGCAGCCCCCCGCAGACGCTCAACATCCTGGTGGATACAGGCAGCAGTAACTTTGCA
V  G  S  P  P  Q  T  L  N  I  L  V  D  T  G  S  S  N  F  A

GTGGGTGCTGCCCCCCACCCCTTCCTGCATCGCTACTACCAGAGGCAGCTGTCCAGCACA
V  G  A  A  P  H  P  F  L  H  R  Y  Y  Q  R  Q  L  S  S  T

TACCGGGACCTCCGGAAGGGTGTGTATGTGCCCTACACCCAGGGCAAGTGGGAAGGGGAG
Y  R  D  L  R  K  G  V  Y  V  P  Y  T  Q  G  K  W  E  G  E

CTGGGCACCGACCTGGTAAGCATCCCCCATGGCCCCAACGTCACTGTGCGTGCCAACATT
L  G  T  D  L  V  S  I  P  H  G  P  N  V  T  V  R  A  N  I

GCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACTGGGAAGGCATCCTG
A  A  I  T  E  S  D  K  F  F  I  N  G  S  N  W  E  G  I  L

GGGCTGGCCTATGCTGAGATTGCCAGGCCTGACGACTCCCTGGAGCCTTTCTTTGACTCT
G  L  A  Y  A  E  I  A  R  P  D  D  S  L  E  P  F  F  D  S

CTGGTAAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGCAGCTTTGTGGTGCTGGCTTC
L  V  K  Q  T  H  V  P  N  L  F  S  L  Q  L  C  G  A  G  F

CCCCTCAACCAGTCTGAAGTGCTGGCCTCTGTCGGAGGGAGCATGATCATTGGAGGTATC
P  L  N  Q  S  E  V  L  A  S  V  G  G  S  M  I  I  G  G  I

GACCACTCGCTGTACACAGGCAGTCTCTGGTATACACCCATCCGGCGGGAGTGGTATTAT
D  H  S  L  Y  T  G  S  L  W  Y  T  P  I  R  R  E  W  Y  Y

GAGGTCATCATTGTGCGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAG
E  V  I  I  V  R  V  E  I  N  G  Q  D  L  K  M  D  C  K  E

TACAACTATGACAAGAGCATTGTGGACAGTGGCACCACCAACCTTCGTTTGCCCAAGAAA
Y  N  Y  D  K  S  I  V  D  S  G  T  T  N  L  R  L  P  K  K

GTGTTTGAAGCTGCAGTCAAATCCATCAAGGCAGCCTCCTCCACGGAGAAGTTCCCTGAT
V  F  E  A  A  V  K  S  I  K  A  A  S  S  T  E  K  F  P  D
```

FIGURE 3 (2)

GGTTTCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGGCACCACCCCTTGGAACATT
G F W L G E Q L V C W Q A G T T P W N I

TTCCCAGTCATCTCACTCTACCTAATGGGTGAGGTTACCAACCAGTCCTTCCGCATCACC
F P V I S L Y L M G E V T N Q S F R I T

ATCCTTCCGCAGCAATACCTGCGGCCAGTGGAAGATGTGGCCACGTCCCAAGACGACTGT
I L P Q Q Y L R P V E D V A T S Q D D C

TACAAGTTTGCCATCTCACAGTCATCCACGGGCACTGTTATGGGAGCTGTTATCATGGAG
Y K F A I S Q S S T G T V M G A V I M E

GGCTTCTACGTTGTCTTTGATCGGGCCCGAAAACGAATTGGCTTTGCTGTCAGCGCTTGC
G F Y V V F D R A R K R I G F A V S A C

CATGTGCACGATGAGTTCAGGACGGCAGCGGTGGAAGGCCCTTTTGTCACCTTGGACATG
H V H D E F R T A A V E G P F V T L D M

GAAGACTGTGGCTACAACATTCCACAGACAGATGAGTCAACCCTCATGACCATAGCCTAT
E D C G Y N I P Q T D E S T L M T I A Y

GTCATGGCTGCCATCTGCGCCCTCTTCATGCTGCCACTCTGCCTCATGGTGTGTCAGTGG
V M A A I C A L F M L P L C L M V C Q W

CGCTGCCTCCGCTGCCTGCGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTG
R C L R C L R Q Q H D D F A D D I S L L

AAGTGAGGAGGCCCATGGGCAGAAGATAGAGATTCCCCTGGACCACACCTCCGTGGTTCA
K

CTTTGGTCACAAGTAGGAGACACAGATGGCACCTGTGGCCAGAGCACCTCAGGACCCTCC
CCACCCACCAAATGCCTCTGCCTTGATGGAGAAGGAAAAGGCTGGCAAGGTGGGTTCCAG
GGACTGTACCTGTAGGAAACAGAAAAGAGAAGAAAGAAGCACTCTGCTGGCGGGAATACT
CTTGGTCACCTCAAATTTAAGTCGGGAAATTCTGCTGCTTGAAACTTCAGCCCTGAACCT
TTGTCCACCATTCCTTTAAATTCTCCAACCCAAAGTATTCTTCTTTTCTTAGTTTCAGAA
GTACTGGCATCACACGCAGGTTACCTTGGCGTGTGTCCCTGTGGTACCCTGGCAGAGAAG
AGACCAAGCTTGTTTCCCTGCTGGCCAAAGTCAGTAGGAGAGGATGCACAGTTTGCTATT
TGCTTTAGAGACAGGGACTGTATAAACAAGCCTAACATTGGTGCAAAGATTGCCTCTTGA
ATTAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 4

```
ATGGCCCCAGCGCTGCACTGGCTCCTGCTATGGGTGGGCTCGGGAATGCTGCCTGCCCAG
M  A  P  A  L  H  W  L  L  L  W  V  G  S  G  M  L  P  A  Q
GGAACCCATCTCGGCATCCGGCTGCCCCTTCGCAGCGGCCTGGCAGGGCCACCCCTGGGC
G  T  H  L  G  I  R  L  P  L  R  S  G  L  A  G  P  P  L  G
CTGAGGCTGCCCCGGGAGACTGACGAGGAATCGGAGGAGCCTGGCCGGAGAGGCAGCTTT
L  R  L  P  R  E  T  D  E  E  S  E  E  P  G  R  R  G  S  F
GTGGAGATGGTGGACAACCTGAGGGGAAAGTCCGGCCAGGGCTACTATGTGGAGATGACC
V  E  M  V  D  N  L  R  G  K  S  G  Q  G  Y  Y  V  E  M  T
GTAGGCAGCCCCCCACAGACGCTCAACATCCTGGTGGACACGGGCAGTAGTAACTTTGCA
V  G  S  P  P  Q  T  L  N  I  L  V  D  T  G  S  S  N  F  A
GTGGGGGCTGCCCCACACCCTTTCCTGCATCGCTACTACCAGAGGCAGCTGTCCAGCACA
V  G  A  A  P  H  P  F  L  H  R  Y  Y  Q  R  Q  L  S  S  T
TATCGAGACCTCCGAAAGGGTGTGTATGTGCCCTACACCCAGGGCAAGTGGGAGGGGGAA
Y  R  D  L  R  K  G  V  Y  V  P  Y  T  Q  G  K  W  E  G  E
CTGGGCACCGACCTGGTGAGCATCCCTCATGGCCCCAACGTCACTGTGCGTGCCAACATT
L  G  T  D  L  V  S  I  P  H  G  P  N  V  T  V  R  A  N  I
GCTGCCATCACTGAATCGGACAAGTTCTTCATCAATGGTTCCAACTGGGAGGGCATCCTA
A  A  I  T  E  S  D  K  F  F  I  N  G  S  N  W  E  G  I  L
GGGCTGGCCTATGCTGAGATTGCCAGGCCCGACGACTCTTTGGAGCCCTTCTTTGACTCC
G  L  A  Y  A  E  I  A  R  P  D  D  S  L  E  P  F  F  D  S
CTGGTGAAGCAGACCCACATTCCCAACATCTTTTCCCTGCAGCTCTGTGGCGCTGGCTTC
L  V  K  Q  T  H  I  P  N  I  F  S  L  Q  L  C  G  A  G  F
CCCCTCAACCAGACCGAGGCACTGGCCTCGGTGGGAGGGAGCATGATCATTGGTGGTATC
P  L  N  Q  T  E  A  L  A  S  V  G  G  S  M  I  I  G  G  I
GACCACTCGCTATACACGGGCAGTCTCTGGTACACACCCATCCGGCGGGAGTGGTATTAT
D  H  S  L  Y  T  G  S  L  W  Y  T  P  I  R  R  E  W  Y  Y
GAAGTGATCATTGTACGTGTGGAAATCAATGGTCAAGATCTCAAGATGGACTGCAAGGAG
E  V  I  I  V  R  V  E  I  N  G  Q  D  L  K  M  D  C  K  E
TACAACTACGACAAGAGCATTGTGGACAGTGGGACCACCAACCTTCGCTTGCCCAAGAAA
Y  N  Y  D  K  S  I  V  D  S  G  T  T  N  L  R  L  P  K  K
GTATTTGAAGCTGCCGTCAAGTCCATCAAGGCAGCCTCCTCGACGGAGAAGTTCCCGGAT
V  F  E  A  A  V  K  S  I  K  A  A  S  S  T  E  K  F  P  D
GGCTTTTGGCTAGGGGAGCAGCTGGTGTGCTGGCAAGCAGGCACGACCCCTTGGAACATT
G  F  W  L  G  E  Q  L  V  C  W  Q  A  G  T  T  P  W  N  I
TTCCCAGTCATTTCACTTTACCTCATGGGTGAAGTCACCAATCAGTCCTTCCGCATCACC
F  P  V  I  S  L  Y  L  M  G  E  V  T  N  Q  S  F  R  I  T
ATCCTTCCTCAGCAATACCTACGGCCGGTGGAGGACGTGGCCACGTCCCAAGACGACTGT
I  L  P  Q  Q  Y  L  R  P  V  E  D  V  A  T  S  Q  D  D  C
TACAAGTTCGCTGTCTCACAGTCATCCACGGGCACTGTTATGGGAGCCGTCATCATGGAA
Y  K  F  A  V  S  Q  S  S  T  G  T  V  M  G  A  V  I  M  E
GGTTTCTATGTCGTCTTCGATCGAGCCCGAAAGCGAATTGGCTTTGCTGTCAGCGCTTGC
G  F  Y  V  V  F  D  R  A  R  K  R  I  G  F  A  V  S  A  C
CATGTGCACGATGAGTTCAGGACGGCGGCAGTGGAAGGTCCGTTTGTTACGGCAGACATG
H  V  H  D  E  F  R  T  A  A  V  E  G  P  F  V  T  A  D  M
GAAGACTGTGGCTACAACATTCCCCAGACAGATGAGTCAACACTTATGACCATAGCCTAT
E  D  C  G  Y  N  I  P  Q  T  D  E  S  T  L  M  T  I  A  Y
GTCATGGCGGCCATCTGCGCCCTCTTCATGTTGCCACTCTGCCTCATGGTATGTCAGTGG
V  M  A  A  I  C  A  L  F  M  L  P  L  C  L  M  V  C  Q  W
CGCTGCCTGCGTTGCCTGCGCCACCAGCACGATGACTTTGCTGATGACATCTCCCTGCTC
R  C  L  R  C  L  R  H  Q  H  D  D  F  A  D  D  I  S  L  L
AAGTAAGGAGGCTCGTGGGCAGATGATGGAGACGCCCCTGGACCACATCTGGGTGGTTCC
K
CTTTGGTCACATGAGTTGGAGCTATGGATGGTACCTGTGGCCAGAGCACCTCAGGACCCT
CACCAACCTGCCAATGCTTCTGGCGTGACAGAACAGAGAAATCAGGCAAGCTGGATTACA
GGGCTTGCACCTGTAGGACACAGGAGAGGGAAGGAAGCAGCGTTCTGGTGGCAGGAATAT
CCTTAGGCACCACAAACTTGAGTTGGAAATTTTGCTGCTTGAAGCTTCAGCCCTGACCCT
CTGCCCAGCATCCTTTAGAGTCTCCAACCTAAAGTATTCTTTATGTCCTTCCAGAAGTAC
TGGCGTCATACTCAGGCTACCCGGCATGTGTCCCTGTGGTACCCTGGCAGAGAAAGGGCC
AATCTCATTCCCTGCTGGCCAAAGTCAGCAGAAGAAGGTGAAGTTTGCCAGTTGCTTTAG
TGATAGGGACTGCAGACTCAAGCCTACACTGGTACAAAGACTGCGTCTTGAGATAAACAA
GAA
```

FIGURE 5

```
  1 MAQALPWLLLWMGAGVLPAHGTQHGIRLPLRSGLGGAPLGLRLPRETDEE 50
    || || |||||.|.|.||| ||  |||||||||| | |||||||||||||
  1 MAPALHWLLLWVGSGMLPAQGTHLGIRLPLRSGLAGPPLGLRLPRETDEE 50
             .         .         .         .         .
 51 PEEPGRRGSFVEMVDNLRGKSGQGYYVEMTVGSPPQTLNILVDTGSSNFA 100
     |||||||||||||||||||||||||||||||||||||||||||||||||
 51 SEEPGRRGSFVEMVDNLRGKSGQGYYVEMTVGSPPQTLNILVDTGSSNFA 100
             .         .         .         .         .
101 VGAAPHPFLHRYYQRQLSSTYRDLRKGVYVPYTQGKWEGELGTDLVSIPH 150
    ||||||||||||||||||||||||||||||||||||||||||||||||||
101 VGAAPHPFLHRYYQRQLSSTYRDLRKGVYVPYTQGKWEGELGTDLVSIPH 150
             .         .         .         .         .
151 GPNVTVRANIAAITESDKFFINGSNWEGILGLAYAEIARPDDSLEPFFDS 200
    ||||||||||||||||||||||||||||||||||||||||||||||||||
151 GPNVTVRANIAAITESDKFFINGSNWEGILGLAYAEIARPDDSLEPFFDS 200
             .         .         .         .         .
201 LVKQTHVPNLFSLQLCGAGFPLNQSEVLASVGGSMIIGGIDHSLYTGSLW 250
    ||||||:||:||||||||||||||.| |||||||||||||||||||||||
201 LVKQTHIPNIFSLQLCGAGFPLNQTEALASVGGSMIIGGIDHSLYTGSLW 250
             .         .         .         .         .
251 YTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRLPKK 300
    ||||||||||||||||||||||||||||||||||||||||||||||||||
251 YTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRLPKK 300
             .         .         .         .         .
301 VFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQAGTTPWNIFPVISLYLMG 350
    ||||||||||||||||||||||||||||||||||||||||||||||||||
301 VFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQAGTTPWNIFPVISLYLMG 350
             .         .         .         .         .
351 EVTNQSFRITILPQQYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIME 400
    ||||||||||||||||||||||||||||||||||:|||||||||||||||
351 EVTNQSFRITILPQQYLRPVEDVATSQDDCYKFAVSQSSTGTVMGAVIME 400
             .         .         .         .         .
401 GFYVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQT 450
    ||||||||||||||||||||||||||||||||||||| ||||||||||||
401 GFYVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTADMEDCGYNIPQT 450
             .         .         .         .         .
451 DESTLMTIAYVMAAICALFMLPLCLMVCQWRCLRCLRQQHDDFADDISLL 500
    ||||||||||||||||||||||||||||||||||||| ||||||||||||
451 DESTLMTIAYVMAAICALFMLPLCLMVCQWRCLRCLRHQHDDFADDISLL 500

501 K 501
    |
501 K 501
```

FIGURE 6 (1)

```
ATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGCGGATCCACCCAGCACGGCATCCGG
M  A  S  M  T  G  G  Q  Q  M  G  R  G  S  T  Q  H  G  I  R

CTGCCCCTGCGCAGCGGCCTGGGGGGCGCCCCCCTGGGGCTGCGGCTGCCCCGGGAGACC
L  P  L  R  S  G  L  G  G  A  P  L  G  L  R  L  P  R  E  T

GACGAAGAGCCCGAGGAGCCCGGCCGGAGGGGCAGCTTTGTGGAGATGGTGGACAACCTG
D  E  E  P  E  E  P  G  R  R  G  S  F  V  E  M  V  D  N  L

AGGGGCAAGTCGGGGCAGGGCTACTACGTGGAGATGACCGTGGGCAGCCCCCCGCAGACG
R  G  K  S  G  Q  G  Y  Y  V  E  M  T  V  G  S  P  P  Q  T

CTCAACATCCTGGTGGATACAGGCAGCAGTAACTTTGCAGTGGGTGCTGCCCCCCACCCC
L  N  I  L  V  D  T  G  S  S  N  F  A  V  G  A  A  P  H  P

TTCCTGCATCGCTACTACCAGAGGCAGCTGTCCAGCACATACCGGGACCTCCGGAAGGGC
F  L  H  R  Y  Y  Q  R  Q  L  S  S  T  Y  R  D  L  R  K  G

GTGTATGTGCCCTACACCCAGGGCAAGTGGGAAGGGGAGCTGGGCACCGACCTGGTAAGC
V  Y  V  P  Y  T  Q  G  K  W  E  G  E  L  G  T  D  L  V  S

ATCCCCCATGGCCCCAACGTCACTGTGCGTGCCAACATTGCTGCCATCACTGAATCAGAC
I  P  H  G  P  N  V  T  V  R  A  N  I  A  A  I  T  E  S  D

AAGTTCTTCATCAACGGCTCCAACTGGGAAGGCATCCTGGGGCTGGCCTATGCTGAGATT
K  F  F  I  N  G  S  N  W  E  G  I  L  G  L  A  Y  A  E  I

GCCAGGCCTGACGACTCCCTGGAGCCTTTCTTTGACTCTCTGGTAAAGCAGACCCACGTT
A  R  P  D  D  S  L  E  P  F  F  D  S  L  V  K  Q  T  H  V

CCCAACCTCTTCTCCCTGCAGCTTTGTGGTGCTGGCTTCCCCCTCAACCAGTCTGAAGTG
P  N  L  F  S  L  Q  L  C  G  A  G  F  P  L  N  Q  S  E  V

CTGGCCTCTGTCGGAGGGAGCATGATCATTGGAGGTATCGACCACTCGCTGTACACAGGC
L  A  S  V  G  G  S  M  I  I  G  G  I  D  H  S  L  Y  T  G

AGTCTCTGGTATACACCCATCCGGCGGGAGTGGTATTATGAGGTCATCATTGTGCGGGTG
S  L  W  Y  T  P  I  R  R  E  W  Y  Y  E  V  I  I  V  R  V

GAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAGTACAACTATGACAAGAGCATT
E  I  N  G  Q  D  L  K  M  D  C  K  E  Y  N  Y  D  K  S  I

GTGGACAGTGGCACCACCAACCTTCGTTTGCCCAAGAAAGTGTTTGAAGCTGCAGTCAAA
V  D  S  G  T  T  N  L  R  L  P  K  K  V  F  E  A  A  V  K

TCCATCAAGGCAGCCTCCTCCACGGAGAAGTTCCCTGATGGTTTCTGGCTAGGAGAGCAG
S  I  K  A  A  S  S  T  E  K  F  P  D  G  F  W  L  G  E  Q

CTGGTGTGCTGGCAAGCAGGCACCACCCCTTGGAACATTTTCCCAGTCATCTCACTCTAC
L  V  C  W  Q  A  G  T  T  P  W  N  I  F  P  V  I  S  L  Y

CTAATGGGTGAGGTTACCAACCAGTCCTTCCGCATCACCATCCTTCCGCAGCAATACCTG
L  M  G  E  V  T  N  Q  S  F  R  I  T  I  L  P  Q  Q  Y  L

CGGCCAGTGGAAGATGTGGCCACGTCCCAAGACGACTGTTACAAGTTTGCCATCTCACAG
```

FIGURE 6 (2)

R P V E D V A T S Q D D C Y K F A I S Q

TCATCCACGGGCACTGTTATGGGAGCTGTTATCATGGAGGGCTTCTACGTTGTCTTTGAT
S S T G T V M G A V I M E G F Y V V F D

CGGGCCCGAAAACGAATTGGCTTTGCTGTCAGCGCTTGCCATGTGCACGATGAGTTCAGG
R A R K R I G F A V S A C H V H D E F R

ACGGCAGCGGTGGAAGGCCCTTTTGTCACCTTGGACATGGAAGACTGTGGCTACAACATT
T A A V E G P F V T L D M E D C G Y N I

CCACAGACAGATGAGTCATGA
P Q T D E S *

FIGURE 7 (1)

ATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGCGGATCGATGACTATCTCTGACTCT
M A S M T G G Q Q M G R G S M T I S D S

CCGCGTGAACAGGACGGATCCACCCAGCACGGCATCCGGCTGCCCCTGCGCAGCGGCCTG
P R E Q D G S T Q H G I R L P L R S G L

GGGGGCGCCCCCCTGGGGCTGCGGCTGCCCCGGGAGACCGACGAAGAGCCCGAGGAGCCC
G G A P L G L R L P R E T D E E P E E P

GGCCGGAGGGGCAGCTTTGTGGAGATGGTGGACAACCTGAGGGGCAAGTCGGGGCAGGGC
G R R G S F V E M V D N L R G K S G Q G

TACTACGTGGAGATGACCGTGGGCAGCCCCCCGCAGACGCTCAACATCCTGGTGGATACA
Y Y V E M T V G S P P Q T L N I L V D T

GGCAGCAGTAACTTTGCAGTGGGTGCTGCCCCCCACCCCTTCCTGCATCGCTACTACCAG
G S S N F A V G A A P H P F L H R Y Y Q

AGGCAGCTGTCCAGCACATACCGGGACCTCCGGAAGGGCGTGTATGTGCCCTACACCCAG
R Q L S S T Y R D L R K G V Y V P Y T Q

GGCAAGTGGGAAGGGGAGCTGGGCACCGACCTGGTAAGCATCCCCCATGGCCCCAACGTC
G K W E G E L G T D L V S I P H G P N V

ACTGTGCGTGCCAACATTGCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCC
T V R A N I A A I T E S D K F F I N G S

AACTGGGAAGGCATCCTGGGGCTGGCCTATGCTGAGATTGCCAGGCCTGACGACTCCCTG
N W E G I L G L A Y A E I A R P D D S L

GAGCCTTTCTTTGACTCTCTGGTAAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGCAG
E P F F D S L V K Q T H V P N L F S L Q

CTTTGTGGTGCTGGCTTCCCCCTCAACCAGTCTGAAGTGCTGGCCTCTGTCGGAGGGAGC
L C G A G F P L N Q S E V L A S V G G S

ATGATCATTGGAGGTATCGACCACTCGCTGTACACAGGCAGTCTCTGGTATACACCCATC
M I I G G I D H S L Y T G S L W Y T P I

CGGCGGGAGTGGTATTATGAGGTCATCATTGTGCGGGTGGAGATCAATGGACAGGATCTG
R R E W Y Y E V I I V R V E I N G Q D L

AAAATGGACTGCAAGGAGTACAACTATGACAAGAGCATTGTGGACAGTGGCACCACCAAC
K M D C K E Y N Y D K S I V D S G T T N

CTTCGTTTGCCCAAGAAAGTGTTTGAAGCTGCAGTCAAATCCATCAAGGCAGCCTCCTCC
L R L P K K V F E A A V K S I K A A S S

ACGGAGAAGTTCCCTGATGGTTTCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGGC
T E K F P D G F W L G E Q L V C W Q A G

ACCACCCCTTGGAACATTTTCCCAGTCATCTCACTCTACCTAATGGGTGAGGTTACCAAC
T T P W N I F P V I S L Y L M G E V T N

FIGURE 7 (2)

```
CAGTCCTTCCGCATCACCATCCTTCCGCAGCAATACCTGCGGCCAGTGGAAGATGTGGCC
Q  S  F  R  I  T  I  L  P  Q  Q  Y  L  R  P  V  E  D  V  A

ACGTCCCAAGACGACTGTTACAAGTTTGCCATCTCACAGTCATCCACGGGCACTGTTATG
T  S  Q  D  D  C  Y  K  F  A  I  S  Q  S  S  T  G  T  V  M

GGAGCTGTTATCATGGAGGGCTTCTACGTTGTCTTTGATCGGGCCCGAAAACGAATTGGC
G  A  V  I  M  E  G  F  Y  V  V  F  D  R  A  R  K  R  I  G

TTTGCTGTCAGCGCTTGCCATGTGCACGATGAGTTCAGGACGGCAGCGGTGGAAGGCCCT
F  A  V  S  A  C  H  V  H  D  E  F  R  T  A  A  V  E  G  P

TTTGTCACCTTGGACATGGAAGACTGTGGCTACAACATTCCACAGACAGATGAGTCATGA
F  V  T  L  D  M  E  D  C  G  Y  N  I  P  Q  T  D  E  S  *
```

FIGURE 8 (1)

```
ATGACTCAGCATGGTATTCGTCTGCCACTGCGTAGCGGTCTGGGTGGTGCTCCACTGGGT
M  T  Q  H  G  I  R  L  P  L  R  S  G  L  G  G  A  P  L  G   -

CTGCGTCTGCCCCGGGAGACCGACGAAGAGCCCGAGGAGCCCGGCCGGAGGGGCAGCTTT
L  R  L  P  R  E  T  D  E  E  P  E  E  P  G  R  R  G  S  F   -

GTGGAGATGGTGGACAACCTGAGGGGCAAGTCGGGGCAGGGCTACTACGTGGAGATGACC
V  E  M  V  D  N  L  R  G  K  S  G  Q  G  Y  Y  V  E  M  T   -

GTGGGCAGCCCCCCGCAGACGCTCAACATCCTGGTGGATACAGGCAGCAGTAACTTTGCA
V  G  S  P  P  Q  T  L  N  I  L  V  D  T  G  S  S  N  F  A   -

GTGGGTGCTGCCCCCCACCCCTTCCTGCATCGCTACTACCAGAGGCAGCTGTCCAGCACA
V  G  A  A  P  H  P  F  L  H  R  Y  Y  Q  R  Q  L  S  S  T   -

TACCGGGACCTCCGGAAGGGCGTGTATGTGCCCTACACCCAGGGCAAGTGGGAAGGGGAG
Y  R  D  L  R  K  G  V  Y  V  P  Y  T  Q  G  K  W  E  G  E   -

CTGGGCACCGACCTGGTAAGCATCCCCCATGGCCCCAACGTCACTGTGCGTGCCAACATT
L  G  T  D  L  V  S  I  P  H  G  P  N  V  T  V  R  A  N  I   -

GCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACTGGGAAGGCATCCTG
A  A  I  T  E  S  D  K  F  F  I  N  G  S  N  W  E  G  I  L   -

GGGCTGGCCTATGCTGAGATTGCCAGGCCTGACGACTCCCTGGAGCCTTTCTTTGACTCT
G  L  A  Y  A  E  I  A  R  P  D  D  S  L  E  P  F  F  D  S

CTGGTAAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGCAGCTTTGTGGTGCTGGCTTC
L  V  K  Q  T  H  V  P  N  L  F  S  L  Q  L  C  G  A  G  F   -

CCCCTCAACCAGTCTGAAGTGCTGGCCTCTGTCGGAGGGAGCATGATCATTGGAGGTATC
P  L  N  Q  S  E  V  L  A  S  V  G  G  S  M  I  I  G  G  I   -

GACCACTCGCTGTACACAGGCAGTCTCTGGTATACACCCATCCGGCGGGAGTGGTATTAT
D  H  S  L  Y  T  G  S  L  W  Y  T  P  I  R  R  E  W  Y  Y   -

GAGGTCATCATTGTGCGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAG
E  V  I  I  V  R  V  E  I  N  G  Q  D  L  K  M  D  C  K  E

TACAACTATGACAAGAGCATTGTGGACAGTGGCACCACCAACCTTCGTTTGCCCAAGAAA
Y  N  Y  D  K  S  I  V  D  S  G  T  T  N  L  R  L  P  K  K   -

GTGTTTGAAGCTGCAGTCAAATCCATCAAGGCAGCCTCCTCCACGGAGAAGTTCCCTGAT
V  F  E  A  A  V  K  S  I  K  A  A  S  S  T  E  K  F  P  D   -

GGTTTCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGGCACCACCCCTTGGAACATT
G  F  W  L  G  E  Q  L  V  C  W  Q  A  G  T  T  P  W  N  I   -

TTCCCAGTCATCTCACTCTACCTAATGGGTGAGGTTACCAACCAGTCCTTTCGCATCACC
F  P  V  I  S  L  Y  L  M  G  E  V  T  N  Q  S  F  R  I  T   -

ATCCTTCCGCAGCAATACCTGCGGCCAGTGGAAGATGTGGCCACGTCCCAAGACGACTGT
I  L  P  Q  Q  Y  L  R  P  V  E  D  V  A  T  S  Q  D  D  C   -
```

FIGURE 8 (2)

```
TACAAGTTTGCCATCTCACAGTCATCCACGGGCACTGTTATGGGAGCTGTTATCATGGAG
Y  K  F  A  I  S  Q  S  S  T  G  T  V  M  G  A  V  I  M  E   -

GGCTTCTACGTTGTCTTTGATCGGGCCCGAAAACGAATTGGCTTTGCTGTCAGCGCTTGC
G  F  Y  V  V  F  D  R  A  R  K  R  I  G  F  A  V  S  A  C   -

CATTAG
H  *
```

FIGURE 9

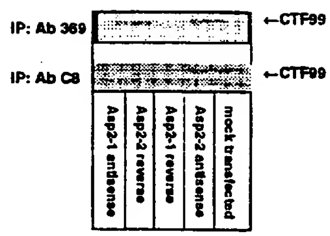
IP: Ab 369
IP: Ab C8
←CTF99
←CTF99
Asp2-1 antisense
Asp2-2 reverse
Asp2-1 reverse
Asp2-2 antisense
mock transfected

FIGURE 10

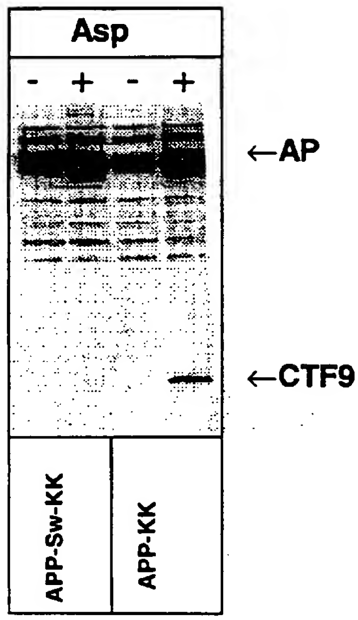

FIGURE 11

<u>MAQALPWLLLWMGAGVLPAHG</u>TQHGIRLPLRSGLGGAPLGLRLPRETDEE
PEEPGRRGSFVEMVDNLRGKSGQGYYVEMTVGSPPQTLNILVDTGSSNFA
VGAAPHPFLHRYYQRQLSSTYRDLRKGVYVPYTQGKWEGELGTDLVSIPH
GPNVTVRANIAAITESDKFFINGSNWEGILGLAYAEIARPDDSLEPFFDS
LVKQTHVPNLFSLQLCGAGFPLNQSEVLASVGGSMIIGGIDHSLYTGSLW
YTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRLPKK
VFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQAGTTPWNIFPVISLYLMG
EVTNQSFRITILPQQYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIME
GFYVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQT
DES

FIGURE 12

MAQALPWLLLWMGAGVLPAHGTQHGIRLPLRSGLGGAPLGLRLPRETDEE
PEEPGRRGSFVEMVDNLRGKSGQGYYVEMTVGSPPQTLNILVDTGSSNFA
VGAAPHPFLHRYYQRQLSSTYRDLRKGVYVPYTQGKWEGELGTDLVSIPH
GPNVTVRANIAAITESDKFFINGSNWEGILGLAYAEIARPDDSLEPFFDS
LVKQTHVPNLFSLQLCGAGFPLNQSEVLASVGGSMIIGGIDHSLYTGSLW
YTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRLPKK
VFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQAGTTPWNIFPVISLYLMG
EVTNQSFRITILPQQYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIME
GFYVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQT
DESHHHHHH

SEQUENCE LISTING

<110> Gurney, Mark E.

Bienkowski, Michael J.

Heinrikson, Robert L.

Parodi, Luis A.

Yan, Riqiang

Pharmacia & Upjohn Company

<120> Alzheimer's Disease Secretase

<130> 6177.P CP

<140>

<141>

<150> 60/101,594

<151> 1998-09-24

<160> 49

<170> PatentIn Ver. 2.0

<210> 1

<211> 1804

<212> DNA

<213> Homo sapiens

<400> 1

atgggcgcac tggcccgggc gctgctgctg cctctgctgg cccagtggct cctgcgcgcc 60

gccccggagc tggcccccgc gcccttcacg ctgcccctcc gggtggccgc ggccacgaac 120

cgcgtagttg cgcccacccc gggacccggg acccctgccg agcgccacgc cgacggcttg 180
gcgctcgccc tggagcctgc cctggcgtcc cccgcgggcg ccgccaactt cttggccatg 240
gtagacaacc tgcagggggga ctctggccgc ggctactacc tggagatgct gatcgggacc 300
cccccgcaga agctacagat tctcgttgac actggaagca gtaactttgc cgtggcagga 360
accccgcact cctacataga cacgtacttt gacacagaga ggtctagcac ataccgctcc 420
aagggctttg acgtcacagt gaagtacaca caaggaagct ggacgggctt cgttggggaa 480
gacctcgtca ccatccccaa aggcttcaat acttcttttc ttgtcaacat tgccactatt 540
tttgaatcag agaatttctt tttgcctggg attaaatgga atggaatact tggcctagct 600
tatgccacac ttgccaagcc atcaagttct ctggagacct tcttcgactc cctggtgaca 660
caagcaaaca tccccaacgt tttctccatg cagatgtgtg gagccggctt gcccgttgct 720
ggatctggga ccaacggagg tagtcttgtc ttgggtggaa ttgaaccaag tttgtataaa 780
ggagacatct ggtatacccc tattaaggaa gagtggtact accagataga aattctgaaa 840
ttggaaattg gaggccaaag ccttaatctg gactgcagag agtataacgc agacaaggcc 900
atcgtggaca gtggcaccac gctgctgcgc ctgccccaga aggtgtttga tgcggtggtg 960
gaagctgtgg cccgcgcatc tctgattcca gaattctctg atggtttctg gactgggtcc 1020
cagctggcgt gctggacgaa ttcggaaaca ccttggtctt acttccctaa aatctccatc 1080
tacctgagag atgagaactc cagcaggtca ttccgtatca caatcctgcc tcagctttac 1140
attcagccca tgatgggggc cggcctgaat tatgaatgtt accgattcgg catttcccca 1200
tccacaaatg cgctggtgat cggtgccacg gtgatggagg gcttctacgt catcttcgac 1260
agagcccaga agagggtggg cttcgcagcg agcccctgtg cagaaattgc aggtgctgca 1320
gtgtctgaaa tttccgggcc tttctcaaca gaggatgtag ccagcaactg tgtccccgct 1380
cagtctttga gcgagcccat tttgtggatt gtgtcctatg cgctcatgag cgtctgtgga 1440
gccatcctcc ttgtcttaat cgtcctgctg ctgctgccgt tccggtgtca gcgtcgcccc 1500
cgtgaccctg aggtcgtcaa tgatgagtcc tctctggtca gacatcgctg gaaatgaata 1560
gccaggcctg acctcaagca accatgaact cagctattaa gaaaatcaca tttccagggc 1620
agcagccggg atcgatggtg gcgctttctc ctgtgcccac ccgtcttcaa tctctgttct 1680
gctcccagat gccttctaga ttcactgtct tttgattctt gattttcaag ctttcaaatc 1740
ctccctactt ccaagaaaaa taattaaaaa aaaaacttca ttctaaacca aaaaaaaaaa 1800
aaaa 1804

<210> 2

<211> 518

<212> PRT

<213> Homo sapiens

<400> 2

Met Gly Ala Leu Ala Arg Ala Leu Leu Leu Pro Leu Leu Ala Gln Trp
1 5 10 15

Leu Leu Arg Ala Ala Pro Glu Leu Ala Pro Ala Pro Phe Thr Leu Pro
20 25 30

Leu Arg Val Ala Ala Ala Thr Asn Arg Val Val Ala Pro Thr Pro Gly
35 40 45

Pro Gly Thr Pro Ala Glu Arg His Ala Asp Gly Leu Ala Leu Ala Leu
50 55 60

Glu Pro Ala Leu Ala Ser Pro Ala Gly Ala Ala Asn Phe Leu Ala Met
65 70 75 80

Val Asp Asn Leu Gln Gly Asp Ser Gly Arg Gly Tyr Tyr Leu Glu Met
85 90 95

Leu Ile Gly Thr Pro Pro Gln Lys Leu Gln Ile Leu Val Asp Thr Gly
100 105 110

Ser Ser Asn Phe Ala Val Ala Gly Thr Pro His Ser Tyr Ile Asp Thr
115 120 125

Tyr Phe Asp Thr Glu Arg Ser Ser Thr Tyr Arg Ser Lys Gly Phe Asp
130 135 140

Val Thr Val Lys Tyr Thr Gln Gly Ser Trp Thr Gly Phe Val Gly Glu
145 150 155 160

Asp Leu Val Thr Ile Pro Lys Gly Phe Asn Thr Ser Phe Leu Val Asn
165 170 175

Ile Ala Thr Ile Phe Glu Ser Glu Asn Phe Phe Leu Pro Gly Ile Lys
180 185 190

Trp Asn Gly Ile Leu Gly Leu Ala Tyr Ala Thr Leu Ala Lys Pro Ser
195 200 205

Ser Ser Leu Glu Thr Phe Phe Asp Ser Leu Val Thr Gln Ala Asn Ile
210 215 220

Pro Asn Val Phe Ser Met Gln Met Cys Gly Ala Gly Leu Pro Val Ala
225 230 235 240

Gly Ser Gly Thr Asn Gly Gly Ser Leu Val Leu Gly Gly Ile Glu Pro
245 250 255

Ser Leu Tyr Lys Gly Asp Ile Trp Tyr Thr Pro Ile Lys Glu Glu Trp
260 265 270

Tyr Tyr Gln Ile Glu Ile Leu Lys Leu Glu Ile Gly Gly Gln Ser Leu
275 280 285

Asn Leu Asp Cys Arg Glu Tyr Asn Ala Asp Lys Ala Ile Val Asp Ser
290 295 300

Gly Thr Thr Leu Leu Arg Leu Pro Gln Lys Val Phe Asp Ala Val Val
305 310 315 320

Glu Ala Val Ala Arg Ala Ser Leu Ile Pro Glu Phe Ser Asp Gly Phe
325 330 335

Trp Thr Gly Ser Gln Leu Ala Cys Trp Thr Asn Ser Glu Thr Pro Trp
340 345 350

Ser Tyr Phe Pro Lys Ile Ser Ile Tyr Leu Arg Asp Glu Asn Ser Ser
355 360 365

Arg Ser Phe Arg Ile Thr Ile Leu Pro Gln Leu Tyr Ile Gln Pro Met
370 375 380

Met Gly Ala Gly Leu Asn Tyr Glu Cys Tyr Arg Phe Gly Ile Ser Pro
385 390 395 400

Ser Thr Asn Ala Leu Val Ile Gly Ala Thr Val Met Glu Gly Phe Tyr
405 410 415

Val Ile Phe Asp Arg Ala Gln Lys Arg Val Gly Phe Ala Ala Ser Pro
420 425 430

Cys Ala Glu Ile Ala Gly Ala Ala Val Ser Glu Ile Ser Gly Pro Phe
435 440 445

Ser Thr Glu Asp Val Ala Ser Asn Cys Val Pro Ala Gln Ser Leu Ser
450 455 460

Glu Pro Ile Leu Trp Ile Val Ser Tyr Ala Leu Met Ser Val Cys Gly

465 470 475 480

Ala Ile Leu Leu Val Leu Ile Val Leu Leu Leu Leu Pro Phe Arg Cys

485 490 495

Gln Arg Arg Pro Arg Asp Pro Glu Val Val Asn Asp Glu Ser Ser Leu

500 505 510

Val Arg His Arg Trp Lys

515

<210> 3

<211> 2070

<212> DNA

<213> Homo sapiens

<400> 3

atggcccaag ccctgccctg gctcctgctg tggatgggcg cgggagtgct gcctgcccac 60

ggcacccagc acggcatccg gctgcccctg cgcagcggcc tgggggggcgc ccccctgggg 120

ctgcggctgc cccgggagac cgacgaagag cccgaggagc ccggccggag ggcagcttt 180

gtggagatgg tggacaacct gaggggcaag tcggggcagg gctactacgt ggagatgacc 240

gtgggcagcc ccccgcagac gctcaacatc ctggtggata caggcagcag taactttgca 300

gtgggtgctg ccccccaccc cttcctgcat cgctactacc agaggcagct gtccagcaca 360

taccgggacc tccggaaggg tgtgtatgtg ccctacaccc agggcaagtg ggaaggggag 420

ctgggcaccg acctggtaag catcccccat ggccccaacg tcactgtgcg tgccaacatt 480

gctgccatca ctgaatcaga caagttcttc atcaacggct ccaactggga aggcatcctg 540

gggctggcct atgctgagat tgccaggcct gacgactccc tggagccttt ctttgactct 600

ctggtaaagc agacccacgt tcccaacctc ttctccctgc acctttgtgg tgctggcttc 660

cccctcaacc agtctgaagt gctggcctct gtcggaggga gcatgatcat tggaggtatc 720

gaccactcgc tgtacacagg cagtctctgg tatacaccca tccggcggga gtggtattat 780

gaggtcatca ttgtgcgggt ggagatcaat ggacaggatc tgaaaatgga ctgcaaggag 840

tacaactatg acaagagcat tgtggacagt ggcaccacca accttcgttt gcccaagaaa 900

gtgtttgaag ctgcagtcaa atccatcaag gcagcctcct ccacggagaa gttccctgat 960

ggtttctggc taggagagca gctggtgtgc tggcaagcag gcaccacccc ttggaacatt 1020

ttcccagtca tctcactcta cctaatgggt gaggttacca accagtcctt ccgcatcacc 1080

atccttccgc agcaatacct gcggccagtg gaagatgtgg ccacgtccca agacgactgt 1140

tacaagtttg ccatctcaca gtcatccacg ggcactgtta tgggagctgt tatcatggag 1200

ggcttctacg ttgtctttga tcgggcccga aaacgaattg gctttgctgt cagcgcttgc 1260

catgtgcacg atgagttcag gacggcagcg gtggaaggcc cttttgtcac cttggacatg 1320

gaagactgtg gctacaacat tccacagaca gatgagtcaa ccctcatgac catagcctat 1380

gtcatggctg ccatctgcgc cctcttcatg ctgccactct gcctcatggt gtgtcagtgg 1440

cgctgcctcc gctgcctgcg ccagcagcat gatgactttg ctgatgacat ctccctgctg 1500

aagtgaggag gcccatgggc agaagataga gattcccctg gaccacacct ccgtggttca 1560

ctttggtcac aagtaggaga cacagatggc acctgtggcc agagcacctc aggaccctcc 1620

ccacccacca aatgcctctg ccttgatgga gaaggaaaag gctggcaagg tgggttccag 1680

ggactgtacc tgtaggaaac agaaaagaga agaaagaagc actctgctgg cgggaatact 1740

cttggtcacc tcaaatttaa gtcgggaaat tctgctgctt gaaacttcag ccctgaacct 1800

ttgtccacca ttcctttaaa ttctccaacc caaagtattc ttcttttctt agtttcagaa 1860

gtactggcat cacacgcagg ttaccttggc gtgtgtccct gtggtaccct ggcagagaag 1920

agaccaagct tgtttccctg ctggccaaag tcagtaggag aggatgcaca gtttgctatt 1980

tgctttagag acagggactg tataaacaag cctaacattg gtgcaaagat tgcctcttga 2040

attaaaaaaa aaaaaaaaaa aaaaaaaaaa 2070

<210> 4

<211> 501

<212> PRT

<213> Homo sapiens

<400> 4

Met Ala Gln Ala Leu Pro Trp Leu Leu Leu Trp Met Gly Ala Gly Val

1 5 10 15

Leu Pro Ala His Gly Thr Gln His Gly Ile Arg Leu Pro Leu Arg Ser
20 25 30

Gly Leu Gly Gly Ala Pro Leu Gly Leu Arg Leu Pro Arg Glu Thr Asp
35 40 45

Glu Glu Pro Glu Glu Pro Gly Arg Arg Gly Ser Phe Val Glu Met Val
50 55 60

Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val Glu Met Thr
65 70 75 80

Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp Thr Gly Ser
85 90 95

Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu His Arg Tyr
100 105 110

Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys Gly Val
115 120 125

Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly Thr Asp
130 135 140

Leu Val Ser Ile Pro His Gly Pro Asn Val Thr Val Arg Ala Asn Ile
145 150 155 160

Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser Asn Trp
165 170 175

Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg Pro Asp Asp

180 185 190

Ser Leu Glu Pro Phe Phe Asp Ser Leu Val Lys Gln Thr His Val Pro

195 200 205

Asn Leu Phe Ser Leu His Leu Cys Gly Ala Gly Phe Pro Leu Asn Gln

210 215 220

Ser Glu Val Leu Ala Ser Val Gly Gly Ser Met Ile Ile Gly Gly Ile

225 230 235 240

Asp His Ser Leu Tyr Thr Gly Ser Leu Trp Tyr Thr Pro Ile Arg Arg

245 250 255

Glu Trp Tyr Tyr Glu Val Ile Ile Val Arg Val Glu Ile Asn Gly Gln

260 265 270

Asp Leu Lys Met Asp Cys Lys Glu Tyr Asn Tyr Asp Lys Ser Ile Val

275 280 285

Asp Ser Gly Thr Thr Asn Leu Arg Leu Pro Lys Lys Val Phe Glu Ala

290 295 300

Ala Val Lys Ser Ile Lys Ala Ala Ser Ser Thr Glu Lys Phe Pro Asp

305 310 315 320

Gly Phe Trp Leu Gly Glu Gln Leu Val Cys Trp Gln Ala Gly Thr Thr

325 330 335

Pro Trp Asn Ile Phe Pro Val Ile Ser Leu Tyr Leu Met Gly Glu Val

340 345 350

Thr Asn Gln Ser Phe Arg Ile Thr Ile Leu Pro Gln Gln Tyr Leu Arg

355 360 365

Pro Val Glu Asp Val Ala Thr Ser Gln Asp Asp Cys Tyr Lys Phe Ala

370 375 380

Ile Ser Gln Ser Ser Thr Gly Thr Val Met Gly Ala Val Ile Met Glu

385 390 395 400

Gly Phe Tyr Val Val Phe Asp Arg Ala Arg Lys Arg Ile Gly Phe Ala

405 410 415

Val Ser Ala Cys His Val His Asp Glu Phe Arg Thr Ala Ala Val Glu

420 425 430

Gly Pro Phe Val Thr Leu Asp Met Glu Asp Cys Gly Tyr Asn Ile Pro

435 440 445

Gln Thr Asp Glu Ser Thr Leu Met Thr Ile Ala Tyr Val Met Ala Ala

450 455 460

Ile Cys Ala Leu Phe Met Leu Pro Leu Cys Leu Met Val Cys Gln Trp

465 470 475 480

Arg Cys Leu Arg Cys Leu Arg Gln Gln His Asp Asp Phe Ala Asp Asp

485 490 495

Ile Ser Leu Leu Lys

500

<210> 5

<211> 1977

<212> DNA

<213> Homo sapiens

<400> 5

atggcccaag ccctgccctg gctcctgctg tggatgggcg cgggagtgct gcctgcccac 60

ggcacccagc acggcatccg gctgcccctg cgcagcggcc tgggggggcgc cccccctgggg 120

ctgcggctgc cccgggagac cgacgaagag cccgaggagc ccggccggag ggcagcttt 180

gtggagatgg tggacaacct gaggggcaag tcggggcagg gctactacgt ggagatgacc 240

gtgggcagcc ccccgcagac gctcaacatc ctggtggata caggcagcag taactttgca 300

gtgggtgctg ccccccaccc cttcctgcat cgctactacc agaggcagct gtccagcaca 360

taccgggacc tccggaaggg tgtgtatgtg ccctacaccc agggcaagtg ggaaggggag 420

ctgggcaccg acctggtaag catcccccat ggccccaacg tcactgtgcg tgccaacatt 480

gctgccatca ctgaatcaga caagttcttc atcaacggct ccaactggga aggcatcctg 540

gggctggcct atgctgagat tgccaggctt tgtggtgctg gcttccccct caaccagtct 600

gaagtgctgg cctctgtcgg agggagcatg atcattggag gtatcgacca ctcgctgtac 660

acaggcagtc tctggtatac acccatccgg cgggagtggt attatgaggt gatcattgtg 720

cgggtggaga tcaatggaca ggatctgaaa atggactgca aggagtacaa ctatgacaag 780

agcattgtgg acagtggcac caccaacctt cgtttgccca agaaagtgtt tgaagctgca 840

gtcaaatcca tcaaggcagc ctcctccacg gagaagttcc ctgatggttt ctggctagga 900

gagcagctgg tgtgctggca agcaggcacc accccttgga acattttccc agtcatctca 960

ctctacctaa tgggtgaggt taccaaccag tccttccgca tcaccatcct tccgcagcaa 1020

tacctgcggc cagtggaaga tgtggccacg tcccaagacg actgttacaa gtttgccatc 1080

tcacagtcat ccacgggcac tgttatggga gctgttatca tggagggctt ctacgttgtc 1140

tttgatcggg cccgaaaacg aattggcttt gctgtcagcg cttgccatgt gcacgatgag 1200

ttcaggacgg cagcggtgga aggccctttt gtcaccttgg acatggaaga ctgtggctac 1260

aacattccac agacagatga gtcaaccctc atgaccatag cctatgtcat ggctgccatc 1320

tgcgccctct tcatgctgcc actctgcctc atggtgtgtc agtggcgctg cctccgctgc 1380

ctgcgccagc agcatgatga ctttgctgat gacatctccc tgctgaagtg aggaggccca 1440

tgggcagaag atagagattc ccctggacca cacctccgtg gttcactttg gtcacaagta 1500

ggagacacag atggcacctg tggccagagc acctcaggac cctccccacc caccaaatgc 1560

ctctgccttg atggagaagg aaaaggctgg caaggtgggt tccagggact gtacctgtag 1620

gaaacagaaa agagaagaaa gaagcactct gctggcggga atactcttgg tcacctcaaa 1680

tttaagtcgg gaaattctgc tgcttgaaac ttcagccctg aacctttgtc caccattcct 1740

ttaaattctc caacccaaag tattcttctt ttcttagttt cagaagtact ggcatcacac 1800

gcaggttacc ttggcgtgtg tccctgtggt accctggcag agaagagacc aagcttgttt 1860

ccctgctggc caaagtcagt aggagaggat gcacagtttg ctatttgctt tagagacagg 1920

gactgtataa acaagcctaa cattggtgca aagattgcct cttgaaaaaa aaaaaaa 1977

<210> 6

<211> 476

<212> PRT

<213> Homo sapiens

<400> 6

Met Ala Gln Ala Leu Pro Trp Leu Leu Leu Trp Met Gly Ala Gly Val
1 5 10 15

Leu Pro Ala His Gly Thr Gln His Gly Ile Arg Leu Pro Leu Arg Ser
20 25 30

Gly Leu Gly Gly Ala Pro Leu Gly Leu Arg Leu Pro Arg Glu Thr Asp
35 40 45

Glu Glu Pro Glu Glu Pro Gly Arg Arg Gly Ser Phe Val Glu Met Val
50 55 60

Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val Glu Met Thr
65 70 75 80

Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp Thr Gly Ser
85 90 95

Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu His Arg Tyr
100 105 110

Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys Gly Val
115 120 125

Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly Thr Asp
130 135 140

Leu Val Ser Ile Pro His Gly Pro Asn Val Thr Val Arg Ala Asn Ile
145 150 155 160

Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser Asn Trp
165 170 175

Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg Leu Cys Gly
180 185 190

Ala Gly Phe Pro Leu Asn Gln Ser Glu Val Leu Ala Ser Val Gly Gly
195 200 205

Ser Met Ile Ile Gly Gly Ile Asp His Ser Leu Tyr Thr Gly Ser Leu
210 215 220

Trp Tyr Thr Pro Ile Arg Arg Glu Trp Tyr Tyr Glu Val Ile Ile Val
225 230 235 240

Arg Val Glu Ile Asn Gly Gln Asp Leu Lys Met Asp Cys Lys Glu Tyr

245 250 255

Asn Tyr Asp Lys Ser Ile Val Asp Ser Gly Thr Thr Asn Leu Arg Leu

260 265 270

Pro Lys Lys Val Phe Glu Ala Ala Val Lys Ser Ile Lys Ala Ala Ser

275 280 285

Ser Thr Glu Lys Phe Pro Asp Gly Phe Trp Leu Gly Glu Gln Leu Val

290 295 300

Cys Trp Gln Ala Gly Thr Thr Pro Trp Asn Ile Phe Pro Val Ile Ser

305 310 315 320

Leu Tyr Leu Met Gly Glu Val Thr Asn Gln Ser Phe Arg Ile Thr Ile

325 330 335

Leu Pro Gln Gln Tyr Leu Arg Pro Val Glu Asp Val Ala Thr Ser Gln

340 345 350

Asp Asp Cys Tyr Lys Phe Ala Ile Ser Gln Ser Ser Thr Gly Thr Val

355 360 365

Met Gly Ala Val Ile Met Glu Gly Phe Tyr Val Val Phe Asp Arg Ala

370 375 380

Arg Lys Arg Ile Gly Phe Ala Val Ser Ala Cys His Val His Asp Glu

385 390 395 400

Phe Arg Thr Ala Ala Val Glu Gly Pro Phe Val Thr Leu Asp Met Glu

405 410 415

Asp Cys Gly Tyr Asn Ile Pro Gln Thr Asp Glu Ser Thr Leu Met Thr

420 425 430

Ile Ala Tyr Val Met Ala Ala Ile Cys Ala Leu Phe Met Leu Pro Leu

435 440 445

Cys Leu Met Val Cys Gln Trp Arg Cys Leu Arg Cys Leu Arg Gln Gln

450 455 460

His Asp Asp Phe Ala Asp Asp Ile Ser Leu Leu Lys

465 470 475

<210> 7

<211> 2043

<212> DNA

<213> Mus musculus

<400> 7

atggccccag cgctgcactg gctcctgcta tgggtgggct cgggaatgct gcctgcccag 60

ggaacccatc tcggcatccg gctgcccctt cgcagcggcc tggcagggcc acccctgggc 120

ctgaggctgc cccgggagac tgacgaggaa tcggaggagc ctggccggag aggcagcttt 180

gtggagatgg tggacaacct gaggggaaag tccggccagg gctactatgt ggagatgacc 240

gtaggcagcc ccccacagac gctcaacatc ctggtggaca cgggcagtag taactttgca 300

gtgggggctg ccccacaccc tttcctgcat cgctactacc agaggcagct gtccagcaca 360

tatcgagacc tccgaaaggg tgtgtatgtg ccctacaccc agggcaagtg ggagggggaa 420

ctgggcaccg acctggtgag catccctcat ggccccaacg tcactgtgcg tgccaacatt 480

gctgccatca ctgaatcgga caagttcttc atcaatggtt ccaactggga gggcatccta 540

gggctggcct atgctgagat tgccaggccc gacgactctt tggagccctt ctttgactcc 600

ctggtgaagc agacccacat tcccaacatc ttttccctgc agctctgtgg cgctggcttc 660

cccctcaacc agaccgaggc actggcctcg gtgggaggga gcatgatcat tggtggtatc 720

gaccactcgc tatacacggg cagtctctgg tacacaccca tccggcggga gtggtattat 780

gaagtgatca ttgtacgtgt ggaaatcaat ggtcaagatc tcaagatgga ctgcaaggag 840

tacaactacg acaagagcat tgtggacagt gggaccacca accttcgctt gcccaagaaa 900

gtatttgaag ctgccgtcaa gtccatcaag gcagcctcct cgacggagaa gttcccggat 960

ggcttttggc taggggagca gctggtgtgc tggcaagcag gcacgacccc ttggaacatt 1020

ttcccagtca tttcactcta cctcatgggt gaagtcacca atcagtcctt ccgcatcacc 1080

atccttcctc agcaatacct acggccggtg gaggacgtgg ccacgtccca agacgactgt 1140

tacaagttcg ctgtctcaca gtcatccacg ggcactgtta tgggagccgt catcatggaa 1200

ggtttctatg tcgtcttcga tcgagcccga aagcgaattg gctttgctgt cagcgcttgc 1260

catgtgcacg atgagttcag gacggcggca gtggaaggtc cgtttgttac ggcagacatg 1320

gaagactgtg gctacaacat tccccagaca gatgagtcaa cacttatgac catagcctat 1380

gtcatggcgg ccatctgcgc cctcttcatg ttgccactct gcctcatggt atgtcagtgg 1440

cgctgcctgc gttgcctgcg ccaccagcac gatgactttg ctgatgacat ctccctgctc 1500

aagtaaggag gctcgtgggc agatgatgga gacgcccctg gaccacatct gggtggttcc 1560

ctttggtcac atgagttgga gctatggatg gtacctgtgg ccagagcacc tcaggaccct 1620

caccaacctg ccaatgcttc tggcgtgaca gaacagagaa atcaggcaag ctggattaca 1680

gggcttgcac ctgtaggaca caggagaggg aaggaagcag cgttctggtg gcaggaatat 1740

ccttaggcac cacaaacttg agttggaaat tttgctgctt gaagcttcag ccctgacccct 1800

ctgcccagca tcctttagag tctccaacct aaagtattct ttatgtcctt ccagaagtac 1860

tggcgtcata ctcaggctac ccggcatgtg tccctgtggt accctggcag agaaagggcc 1920

aatctcattc cctgctggcc aaagtcagca gaagaaggtg aagtttgcca gttgctttag 1980

tgatagggac tgcagactca agcctacact ggtacaaaga ctgcgtcttg agataaacaa 2040

gaa 2043

<210> 8

<211> 501

<212> PRT

<213> Mus musculus

<400> 8

Met Ala Pro Ala Leu His Trp Leu Leu Leu Trp Val Gly Ser Gly Met
1   5   10   15

Leu Pro Ala Gln Gly Thr His Leu Gly Ile Arg Leu Pro Leu Arg Ser
            20   25   30

Gly Leu Ala Gly Pro Pro Leu Gly Leu Arg Leu Pro Arg Glu Thr Asp
        35   40   45

Glu Glu Ser Glu Glu Pro Gly Arg Arg Gly Ser Phe Val Glu Met Val
    50   55   60

Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val Glu Met Thr
65   70   75   80

Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp Thr Gly Ser
                85   90   95

Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu His Arg Tyr
            100   105   110

Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys Gly Val
        115   120   125

Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly Thr Asp
    130   135   140

Leu Val Ser Ile Pro His Gly Pro Asn Val Thr Val Arg Ala Asn Ile
145   150   155   160

Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser Asn Trp

165 170 175

Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg Pro Asp Asp

180 185 190

Ser Leu Glu Pro Phe Phe Asp Ser Leu Val Lys Gln Thr His Ile Pro

195 200 205

Asn Ile Phe Ser Leu Gln Leu Cys Gly Ala Gly Phe Pro Leu Asn Gln

210 215 220

Thr Glu Ala Leu Ala Ser Val Gly Gly Ser Met Ile Ile Gly Gly Ile

225 230 235 240

Asp His Ser Leu Tyr Thr Gly Ser Leu Trp Tyr Thr Pro Ile Arg Arg

245 250 255

Glu Trp Tyr Tyr Glu Val Ile Ile Val Arg Val Glu Ile Asn Gly Gln

260 265 270

Asp Leu Lys Met Asp Cys Lys Glu Tyr Asn Tyr Asp Lys Ser Ile Val

275 280 285

Asp Ser Gly Thr Thr Asn Leu Arg Leu Pro Lys Lys Val Phe Glu Ala

290 295 300

Ala Val Lys Ser Ile Lys Ala Ala Ser Ser Thr Glu Lys Phe Pro Asp

305 310 315 320

Gly Phe Trp Leu Gly Glu Gln Leu Val Cys Trp Gln Ala Gly Thr Thr

325 330 335

Pro Trp Asn Ile Phe Pro Val Ile Ser Leu Tyr Leu Met Gly Glu Val

340 345 350

Thr Asn Gln Ser Phe Arg Ile Thr Ile Leu Pro Gln Gln Tyr Leu Arg

355 360 365

Pro Val Glu Asp Val Ala Thr Ser Gln Asp Asp Cys Tyr Lys Phe Ala

370 375 380

Val Ser Gln Ser Ser Thr Gly Thr Val Met Gly Ala Val Ile Met Glu

385 390 395 400

Gly Phe Tyr Val Val Phe Asp Arg Ala Arg Lys Arg Ile Gly Phe Ala

405 410 415

Val Ser Ala Cys His Val His Asp Glu Phe Arg Thr Ala Ala Val Glu

420 425 430

Gly Pro Phe Val Thr Ala Asp Met Glu Asp Cys Gly Tyr Asn Ile Pro

435 440 445

Gln Thr Asp Glu Ser Thr Leu Met Thr Ile Ala Tyr Val Met Ala Ala

450 455 460

Ile Cys Ala Leu Phe Met Leu Pro Leu Cys Leu Met Val Cys Gln Trp

465 470 475 480

Arg Cys Leu Arg Cys Leu Arg His Gln His Asp Asp Phe Ala Asp Asp

485 490 495

Ile Ser Leu Leu Lys

500

<210> 9

<211> 2088

<212> DNA

<213> Homo sapiens

<400> 9

atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60

cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120

ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180

acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240

cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300

ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360

gagtttgtaa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420

atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480

aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540

ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600

gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660

agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720

gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780

ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840

gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900

gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960

gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020

gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080

caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140

acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200

tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260

aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320

cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380

gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440

gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500

gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560

tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620

gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680

gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740

tctgggttga caaatatcaa gacggaggag atctctgaag tgaagatgga tgcagaattc 1800

cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860

ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920

atcgtcatca ccttggtgat gctgaagaag aaacagtaca catccattca tcatggtgtg 1980

gtggaggttg acgccgctgt caccccagag gagcgccacc tgtccaagat gcagcagaac 2040

ggctacgaaa atccaaccta caagttcttt gagcagatgc agaactag 2088

<210> 10

<211> 695

<212> PRT

<213> Homo sapiens

<400> 10

Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
1 5 10 15

Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
20 25 30

Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
35 40 45

Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
50 55 60

Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65 70 75 80

Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
85 90 95

Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
100 105 110

Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu
115 120 125

Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
130 135 140

Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145 150 155 160

Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
165 170 175

Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
180 185 190

Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
195 200 205

Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys

210 215 220

Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu

225 230 235 240

Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu

245 250 255

Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile

260 265 270

Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg

275 280 285

Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu

290 295 300

Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys

305 310 315 320

Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg

325 330 335

Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp

340 345 350

Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu

355 360 365

Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala

370 375 380

Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
385 390 395 400

Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
405 410 415

Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
420 425 430

Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
435 440 445

Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
450 455 460

Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
465 470 475 480

Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
485 490 495

Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
500 505 510

Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
515 520 525

Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
530 535 540

Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
545 550 555 560

Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
565 570 575

Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
580 585 590

Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
595 600 605

His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
610 615 620

Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
625 630 635 640

Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile
645 650 655

His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu Arg
660 665 670

His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr Lys
675 680 685

Phe Phe Glu Gln Met Gln Asn
690 695

<210> 11

<211> 2088

<212> DNA

<213> Homo sapiens

<400> 11

atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60

cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120

ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180

acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240

cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300

ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360

gagtttgtaa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420

atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480

aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540

ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600

gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660

agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720

gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780

ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840

gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900

gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960

gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020

gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080

caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140

acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200

tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260

aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320

cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380

gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440

gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500

gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560

tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620

gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680

gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740

tctgggttga caaatatcaa gacggaggag atctctgaag tgaatctgga tgcagaattc 1800

cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860

ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920

atcgtcatca ccttggtgat gctgaagaag aaacagtaca catccattca tcatggtgtg 1980

gtggaggttg acgccgctgt caccccagag gagcgccacc tgtccaagat gcagcagaac 2040

ggctacgaaa atccaaccta caagttcttt gagcagatgc agaactag 2088

<210> 12

<211> 695

<212> PRT

<213> Homo sapiens

<400> 12

Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
1 5 10 15

Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
20 25 30

Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
35 40 45

Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
50 55 60

Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65 70 75 80

Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
85 90 95

Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
100 105 110

Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu
115 120 125

Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
130 135 140

Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145 150 155 160

Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
165 170 175

Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
180 185 190

Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
195 200 205

Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
210 215 220

Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
225 230 235 240

Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu

245 250 255

Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile

260 265 270

Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg

275 280 285

Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu

290 295 300

Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys

305 310 315 320

Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg

325 330 335

Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp

340 345 350

Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu

355 360 365

Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala

370 375 380

Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn

385 390 395 400

Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe

405 410 415

Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His

420 425 430

Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala

435 440 445

Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu

450 455 460

Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala

465 470 475 480

Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn

485 490 495

Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser

500 505 510

Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr

515 520 525

Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln

530 535 540

Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn

545 550 555 560

Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr

565 570 575

Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
580 585 590

Glu Val Asn Leu Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
595 600 605

His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
610 615 620

Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
625 630 635 640

Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile
645 650 655

His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu Arg
660 665 670

His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr Lys
675 680 685

Phe Phe Glu Gln Met Gln Asn
690 695

<210> 13

<211> 2088

<212> DNA

<213> Homo sapiens

<400> 13

atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360
gagtttgtaa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaagatgga tgcagaattc 1800

cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860

ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920

atcttcatca ccttggtgat gctgaagaag aaacagtaca catccattca tcatggtgtg 1980

gtggaggttg acgccgctgt caccccagag gagcgccacc tgtccaagat gcagcagaac 2040

ggctacgaaa atccaaccta caagttcttt gagcagatgc agaactag 2088

<210> 14

<211> 695

<212> PRT

<213> Homo sapiens

<400> 14

Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
1 5 10 15

Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
20 25 30

Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
35 40 45

Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
50 55 60

Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65 70 75 80

Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
85 90 95

Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val

100 105 110

Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu

115 120 125

Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys

130 135 140

Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu

145 150 155 160

Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile

165 170 175

Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu

180 185 190

Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val

195 200 205

Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys

210 215 220

Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu

225 230 235 240

Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu

245 250 255

Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile

260 265 270

Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
275 280 285

Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
290 295 300

Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
305 310 315 320

Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
325 330 335

Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
340 345 350

Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu
355 360 365

Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
370 375 380

Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
385 390 395 400

Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
405 410 415

Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
420 425 430

Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala

435 440 445

Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu

450 455 460

Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala

465 470 475 480

Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn

485 490 495

Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser

500 505 510

Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr

515 520 525

Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln

530 535 540

Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn

545 550 555 560

Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr

565 570 575

Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser

580 585 590

Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val

595 600 605

His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys

610 615 620

Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val

625 630 635 640

Ile Phe Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile

645 650 655

His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu Arg

660 665 670

His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr Lys

675 680 685

Phe Phe Glu Gln Met Gln Asn

690 695

<210> 15

<211> 2094

<212> DNA

<213> Homo sapiens

<400> 15

atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60

cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120

ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180

acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240

cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300

ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360

gagtttgtaa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420

atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480

aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540

ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600

gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660

agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720

gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780

ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840

gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900

gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960

gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020

gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080

caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140

acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200

tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260

aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320

cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380

gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440

gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500

gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560

tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620

gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680

gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740

tctgggttga caaatatcaa gacggaggag atctctgaag tgaagatgga tgcagaattc 1800

cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860

ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920

atcgtcatca ccttggtgat gctgaagaag aaacagtaca catccattca tcatggtgtg 1980

gtggaggttg acgccgctgt caccccagag gagcgccacc tgtccaagat gcagcagaac 2040

ggctacgaaa atccaaccta caagttcttt gagcagatgc agaacaagaa gtag 2094

<210> 16

<211> 697

<212> PRT

<213> Homo sapiens

<400> 16

Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
1 5 10 15

Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
20 25 30

Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
35 40 45

Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
50 55 60

Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65 70 75 80

Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
85 90 95

Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
100 105 110

Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu
115 120 125

Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
130 135 140

Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145 150 155 160

Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
165 170 175

Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
180 185 190

Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
195 200 205

Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
210 215 220

Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
225 230 235 240

Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
245 250 255

Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
260 265 270

Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
275 280 285

Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu

290 295 300

Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys

305 310 315 320

Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg

325 330 335

Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp

340 345 350

Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu

355 360 365

Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala

370 375 380

Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn

385 390 395 400

Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe

405 410 415

Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His

420 425 430

Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala

435 440 445

Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu

450 455 460

Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
465 470 475 480

Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
485 490 495

Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
500 505 510

Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
515 520 525

Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
530 535 540

Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
545 550 555 560

Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
565 570 575

Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
580 585 590

Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
595 600 605

His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
610 615 620

Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
625 630 635 640

Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile
645 650 655

His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu Arg
660 665 670

His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr Lys
675 680 685

Phe Phe Glu Gln Met Gln Asn Lys Lys
690 695

<210> 17

<211> 2094

<212> DNA

<213> Homo sapiens

<400> 17

atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360
gagtttgtaa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540

ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaatctgga tgcagaattc 1800
cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920
atcgtcatca ccttggtgat gctgaagaag aaacagtaca catccattca tcatggtgtg 1980
gtggaggttg acgccgctgt caccccagag gagcgccacc tgtccaagat gcagcagaac 2040
ggctacgaaa atccaaccta caagttcttt gagcagatgc agaacaagaa gtag 2094

<210> 18

<211> 697

<212> PRT

<213> Homo sapiens

<400> 18

Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
1   5   10   15

Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
            20   25   30

Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
        35   40   45

Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
    50   55   60

Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65   70   75   80

Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
                85   90   95

Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
            100   105   110

Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu
        115   120   125

Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
    130   135   140

Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145   150   155   160

Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
165 170 175

Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
180 185 190

Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
195 200 205

Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
210 215 220

Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
225 230 235 240

Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
245 250 255

Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
260 265 270

Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
275 280 285

Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
290 295 300

Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
305 310 315 320

Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg

325 330 335

Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp

340 345 350

Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu

355 360 365

Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala

370 375 380

Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn

385 390 395 400

Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe

405 410 415

Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His

420 425 430

Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala

435 440 445

Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu

450 455 460

Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala

465 470 475 480

Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn

485 490 495

Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser

500 505 510

Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr

515 520 525

Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln

530 535 540

Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn

545 550 555 560

Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr

565 570 575

Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser

580 585 590

Glu Val Asn Leu Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val

595 600 605

His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys

610 615 620

Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val

625 630 635 640

Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile

645 650 655

His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu Arg

660 665 670

His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr Lys

675 680 685

Phe Phe Glu Gln Met Gln Asn Lys Lys

690 695

<210> 19

<211> 2094

<212> DNA

<213> Homo sapiens

<400> 19

atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60

cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120

ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180

acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240

cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300

ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360

gagtttgtaa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420

atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480

aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540

ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600

gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660

agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720

gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780

ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840

gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaagatgga tgcagaattc 1800
cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920
atcttcatca ccttggtgat gctgaagaag aaacagtaca catccattca tcatggtgtg 1980
gtggaggttg acgccgctgt caccccagag gagcgccacc tgtccaagat gcagcagaac 2040
ggctacgaaa atccaaccta caagttcttt gagcagatgc agaacaagaa gtag 2094

<210> 20

<211> 697

<212> PRT

<213> Homo sapiens

<400> 20

Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
1 5 10 15

Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro

20 25 30

Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln

35 40 45

Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp

50 55 60

Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu

65 70 75 80

Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn

85 90 95

Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val

100 105 110

Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu

115 120 125

Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys

130 135 140

Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu

145 150 155 160

Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile

165 170 175

Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu

180 185 190

Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val

195 200 205

Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys

210 215 220

Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu

225 230 235 240

Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu

245 250 255

Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile

260 265 270

Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg

275 280 285

Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu

290 295 300

Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys

305 310 315 320

Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg

325 330 335

Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp

340 345 350

Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu

355 360 365

Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala

370 375 380

Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn

385 390 395 400

Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe

405 410 415

Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His

420 425 430

Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala

435 440 445

Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu

450 455 460

Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala

465 470 475 480

Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn

485 490 495

Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser

500 505 510

Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr

515 520 525

Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln

530 535 540

Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn

545 550 555 560

Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr

565 570 575

Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser

580 585 590

Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val

595 600 605

His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys

610 615 620

Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val

625 630 635 640

Ile Phe Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile

645 650 655

His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu Arg

660 665 670

His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr Lys

675 680 685

Phe Phe Glu Gln Met Gln Asn Lys Lys

690 695

<210> 21

<211> 1341

<212> DNA

<213> Homo sapiens

<400> 21

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccacccagca cggcatccgg 60

ctgcccctgc gcagcggcct gggggcgcc cccctggggc tgcggctgcc ccgggagacc 120

gacgaagagc ccgaggagcc cggccggagg ggcagctttg tggagatggt ggacaacctg 180

aggggcaagt cggggcaggg ctactacgtg gagatgaccg tgggcagccc cccgcagacg 240

ctcaacatcc tggtggatac aggcagcagt aactttgcag tgggtgctgc cccccacccc 300

ttcctgcatc gctactacca gaggcagctg tccagcacat accgggacct ccggaagggt 360

gtgtatgtgc cctacaccca gggcaagtgg gaaggggagc tgggcaccga cctggtaagc 420

atcccccatg gccccaacgt cactgtgcgt gccaacattg ctgccatcac tgaatcagac 480

aagttcttca tcaacggctc caactgggaa ggcatcctgg ggctggccta tgctgagatt 540

gccaggcctg acgactccct ggagcctttc tttgactctc tggtaaagca gacccacgtt 600

cccaacctct tctccctgca cctttgtggt gctggcttcc ccctcaacca gtctgaagtg 660

ctggcctctg tcggagggag catgatcatt ggaggtatcg accactcgct gtacacaggc 720

agtctctggt atacacccat ccggcgggag tggtattatg aggtcatcat tgtgcgggtg 780

gagatcaatg gacaggatct gaaaatggac tgcaaggagt acaactatga caagagcatt 840

gtggacagtg gcaccaccaa ccttcgtttg cccaagaaag tgtttgaagc tgcagtcaaa 900

tccatcaagg cagcctcctc cacggagaag ttccctgatg gtttctggct aggagagcag 960

ctggtgtgct ggcaagcagg caccacccct tggaacattt tcccagtcat ctcactctac 1020

ctaatgggtg aggttaccaa ccagtccttc cgcatcacca tccttccgca gcaatacctg 1080

cggccagtgg aagatgtggc cacgtcccaa gacgactgtt acaagtttgc catctcacag 1140

tcatccacgg gcactgttat gggagctgtt atcatggagg gcttctacgt tgtctttgat 1200

cgggcccgaa aacgaattgg ctttgctgtc agcgcttgcc atgtgcacga tgagttcagg 1260

acggcagcgg tggaaggccc ttttgtcacc ttggacatgg aagactgtgg ctacaacatt 1320

ccacagacag atgagtcatg a 1341

<210> 22

<211> 446

<212> PRT

<213> Homo sapiens

<400> 22

Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Gly Ser Thr Gln
1 5 10 15

His Gly Ile Arg Leu Pro Leu Arg Ser Gly Leu Gly Gly Ala Pro Leu
20 25 30

Gly Leu Arg Leu Pro Arg Glu Thr Asp Glu Glu Pro Glu Glu Pro Gly
35 40 45

Arg Arg Gly Ser Phe Val Glu Met Val Asp Asn Leu Arg Gly Lys Ser
50 55 60

Gly Gln Gly Tyr Tyr Val Glu Met Thr Val Gly Ser Pro Pro Gln Thr
65 70 75 80

Leu Asn Ile Leu Val Asp Thr Gly Ser Ser Asn Phe Ala Val Gly Ala
85 90 95

Ala Pro His Pro Phe Leu His Arg Tyr Tyr Gln Arg Gln Leu Ser Ser
100 105 110

Thr Tyr Arg Asp Leu Arg Lys Gly Val Tyr Val Pro Tyr Thr Gln Gly
115 120 125

Lys Trp Glu Gly Glu Leu Gly Thr Asp Leu Val Ser Ile Pro His Gly
130 135 140

Pro Asn Val Thr Val Arg Ala Asn Ile Ala Ala Ile Thr Glu Ser Asp
145 150 155 160

Lys Phe Phe Ile Asn Gly Ser Asn Trp Glu Gly Ile Leu Gly Leu Ala
165 170 175

Tyr Ala Glu Ile Ala Arg Pro Asp Asp Ser Leu Glu Pro Phe Phe Asp
180 185 190

Ser Leu Val Lys Gln Thr His Val Pro Asn Leu Phe Ser Leu His Leu
195 200 205

Cys Gly Ala Gly Phe Pro Leu Asn Gln Ser Glu Val Leu Ala Ser Val
210 215 220

Gly Gly Ser Met Ile Ile Gly Gly Ile Asp His Ser Leu Tyr Thr Gly
225 230 235 240

Ser Leu Trp Tyr Thr Pro Ile Arg Arg Glu Trp Tyr Tyr Glu Val Ile
245 250 255

Ile Val Arg Val Glu Ile Asn Gly Gln Asp Leu Lys Met Asp Cys Lys
260 265 270

Glu Tyr Asn Tyr Asp Lys Ser Ile Val Asp Ser Gly Thr Thr Asn Leu

275 280 285

Arg Leu Pro Lys Lys Val Phe Glu Ala Ala Val Lys Ser Ile Lys Ala

290 295 300

Ala Ser Ser Thr Glu Lys Phe Pro Asp Gly Phe Trp Leu Gly Glu Gln

305 310 315 320

Leu Val Cys Trp Gln Ala Gly Thr Thr Pro Trp Asn Ile Phe Pro Val

325 330 335

Ile Ser Leu Tyr Leu Met Gly Glu Val Thr Asn Gln Ser Phe Arg Ile

340 345 350

Thr Ile Leu Pro Gln Gln Tyr Leu Arg Pro Val Glu Asp Val Ala Thr

355 360 365

Ser Gln Asp Asp Cys Tyr Lys Phe Ala Ile Ser Gln Ser Ser Thr Gly

370 375 380

Thr Val Met Gly Ala Val Ile Met Glu Gly Phe Tyr Val Val Phe Asp

385 390 395 400

Arg Ala Arg Lys Arg Ile Gly Phe Ala Val Ser Ala Cys His Val His

405 410 415

Asp Glu Phe Arg Thr Ala Ala Val Glu Gly Pro Phe Val Thr Leu Asp

420 425 430

Met Glu Asp Cys Gly Tyr Asn Ile Pro Gln Thr Asp Glu Ser

435 440 445

<210> 23

<211> 1380

<212> DNA

<213> Homo sapiens

<400> 23

atggctagca tgactggtgg acagcaaatg ggtcgcggat cgatgactat ctctgactct 60
ccgcgtgaac aggacggatc cacccagcac ggcatccggc tgcccctgcg cagcggcctg 120
gggggcgccc ccctggggct gcggctgccc cgggagaccg acgaagagcc cgaggagccc 180
ggccggaggg gcagctttgt ggagatggtg gacaacctga ggggcaagtc ggggcagggc 240
tactacgtgg agatgaccgt gggcagcccc ccgcagacgc tcaacatcct ggtggataca 300
ggcagcagta actttgcagt gggtgctgcc ccccacccct tcctgcatcg ctactaccag 360
aggcagctgt ccagcacata ccgggacctc cggaagggtg tgtatgtgcc ctacacccag 420
ggcaagtggg aaggggagct gggcaccgac ctggtaagca tcccccatgg ccccaacgtc 480
actgtgcgtg ccaacattgc tgccatcact gaatcagaca agttcttcat caacggctcc 540
aactgggaag gcatcctggg gctggcctat gctgagattg ccaggcctga cgactccctg 600
gagccttttct ttgactctct ggtaaagcag acccacgttc ccaacctctt ctccctgcac 660
ctttgtggtg ctggcttccc cctcaaccag tctgaagtgc tggcctctgt cggagggagc 720
atgatcattg gaggtatcga ccactcgctg tacacaggca gtctctggta tacacccatc 780
cggcgggagt ggtattatga ggtcatcatt gtgcgggtgg agatcaatgg acaggatctg 840
aaaatggact gcaaggagta caactatgac aagagcattg tggacagtgg caccaccaac 900
cttcgtttgc ccaagaaagt gtttgaagct gcagtcaaat ccatcaaggc agcctcctcc 960
acggagaagt tccctgatgg tttctggcta ggagagcagc tggtgtgctg gcaagcaggc 1020
accacccctt ggaacatttt cccagtcatc tcactctacc taatgggtga ggttaccaac 1080
cagtccttcc gcatcaccat ccttccgcag caatacctgc ggccagtgga agatgtggcc 1140
acgtcccaag acgactgtta caagtttgcc atctcacagt catccacggg cactgttatg 1200
ggagctgtta tcatggaggg cttctacgtt gtctttgatc gggcccgaaa acgaattggc 1260
tttgctgtca gcgcttgcca tgtgcacgat gagttcagga cggcagcggt ggaaggccct 1320

ttgtcacct tggacatgga agactgtggc tacaacattc cacagacaga tgagtcatga 1380

<210> 24

<211> 459

<212> PRT

<213> Homo sapiens

<400> 24

Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Gly Ser Met Thr
1 5 10 15

Ile Ser Asp Ser Pro Arg Glu Gln Asp Gly Ser Thr Gln His Gly Ile
20 25 30

Arg Leu Pro Leu Arg Ser Gly Leu Gly Gly Ala Pro Leu Gly Leu Arg
35 40 45

Leu Pro Arg Glu Thr Asp Glu Glu Pro Glu Glu Pro Gly Arg Arg Gly
50 55 60

Ser Phe Val Glu Met Val Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly
65 70 75 80

Tyr Tyr Val Glu Met Thr Val Gly Ser Pro Pro Gln Thr Leu Asn Ile
85 90 95

Leu Val Asp Thr Gly Ser Ser Asn Phe Ala Val Gly Ala Ala Pro His
100 105 110

Pro Phe Leu His Arg Tyr Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg
115 120 125

Asp Leu Arg Lys Gly Val Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu
130 135 140

Gly Glu Leu Gly Thr Asp Leu Val Ser Ile Pro His Gly Pro Asn Val
145 150 155 160

Thr Val Arg Ala Asn Ile Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe
165 170 175

Ile Asn Gly Ser Asn Trp Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu
180 185 190

Ile Ala Arg Pro Asp Asp Ser Leu Glu Pro Phe Phe Asp Ser Leu Val
195 200 205

Lys Gln Thr His Val Pro Asn Leu Phe Ser Leu His Leu Cys Gly Ala
210 215 220

Gly Phe Pro Leu Asn Gln Ser Glu Val Leu Ala Ser Val Gly Gly Ser
225 230 235 240

Met Ile Ile Gly Gly Ile Asp His Ser Leu Tyr Thr Gly Ser Leu Trp
245 250 255

Tyr Thr Pro Ile Arg Arg Glu Trp Tyr Tyr Glu Val Ile Ile Val Arg
260 265 270

Val Glu Ile Asn Gly Gln Asp Leu Lys Met Asp Cys Lys Glu Tyr Asn
275 280 285

Tyr Asp Lys Ser Ile Val Asp Ser Gly Thr Thr Asn Leu Arg Leu Pro

290 295 300

Lys Lys Val Phe Glu Ala Ala Val Lys Ser Ile Lys Ala Ala Ser Ser

305 310 315 320

Thr Glu Lys Phe Pro Asp Gly Phe Trp Leu Gly Glu Gln Leu Val Cys

325 330 335

Trp Gln Ala Gly Thr Thr Pro Trp Asn Ile Phe Pro Val Ile Ser Leu

340 345 350

Tyr Leu Met Gly Glu Val Thr Asn Gln Ser Phe Arg Ile Thr Ile Leu

355 360 365

Pro Gln Gln Tyr Leu Arg Pro Val Glu Asp Val Ala Thr Ser Gln Asp

370 375 380

Asp Cys Tyr Lys Phe Ala Ile Ser Gln Ser Ser Thr Gly Thr Val Met

385 390 395 400

Gly Ala Val Ile Met Glu Gly Phe Tyr Val Val Phe Asp Arg Ala Arg

405 410 415

Lys Arg Ile Gly Phe Ala Val Ser Ala Cys His Val His Asp Glu Phe

420 425 430

Arg Thr Ala Ala Val Glu Gly Pro Phe Val Thr Leu Asp Met Glu Asp

435 440 445

Cys Gly Tyr Asn Ile Pro Gln Thr Asp Glu Ser

450 455

<210> 25

<211> 1302

<212> DNA

<213> Homo sapiens

<400> 25

atgactcagc atggtattcg tctgccactg cgtagcggtc tgggtggtgc tccactgggt 60

ctgcgtctgc cccgggagac cgacgaagag cccgaggagc ccggccggag gggcagcttt 120

gtggagatgg tggacaacct gaggggcaag tcggggcagg gctactacgt ggagatgacc 180

gtgggcagcc ccccgcagac gctcaacatc ctggtggata caggcagcag taactttgca 240

gtgggtgctg ccccccaccc cttcctgcat cgctactacc agaggcagct gtccagcaca 300

taccgggacc tccggaaggg tgtgtatgtg ccctacaccc agggcaagtg ggaaggggag 360

ctgggcaccg acctggtaag catcccccat ggccccaacg tcactgtgcg tgccaacatt 420

gctgccatca ctgaatcaga caagttcttc atcaacggct ccaactggga aggcatcctg 480

gggctggcct atgctgagat tgccaggcct gacgactccc tggagccttt ctttgactct 540

ctggtaaagc agacccacgt tcccaacctc ttctccctgc acctttgtgg tgctggcttc 600

cccctcaacc agtctgaagt gctggcctct gtcggaggga gcatgatcat tggaggtatc 660

gaccactcgc tgtacacagg cagtctctgg tatacaccca tccggcggga gtggtattat 720

gaggtcatca ttgtgcgggt ggagatcaat ggacaggatc tgaaaatgga ctgcaaggag 780

tacaactatg acaagagcat tgtggacagt ggcaccacca accttcgttt gcccaagaaa 840

gtgtttgaag ctgcagtcaa atccatcaag gcagcctcct ccacggagaa gttccctgat 900

ggtttctggc taggagagca gctggtgtgc tggcaagcag gcaccacccc ttggaacatt 960

ttcccagtca tctcactcta cctaatgggt gaggttacca accagtcctt ccgcatcacc 1020

atccttccgc agcaatacct gcggccagtg gaagatgtgg ccacgtccca agacgactgt 1080

tacaagtttg ccatctcaca gtcatccacg ggcactgtta tgggagctgt tatcatggag 1140

ggcttctacg ttgtctttga tcgggcccga aaacgaattg gctttgctgt cagcgcttgc 1200

catgtgcacg atgagttcag gacggcagcg gtggaaggcc cttttgtcac cttggacatg 1260

gaagactgtg gctacaacat tccacagaca gatgagtcat ga 1302

<210> 26

<211> 433

<212> PRT

<213> Homo sapiens

<400> 26

Met Thr Gln His Gly Ile Arg Leu Pro Leu Arg Ser Gly Leu Gly Gly
1 5 10 15

Ala Pro Leu Gly Leu Arg Leu Pro Arg Glu Thr Asp Glu Glu Pro Glu
20 25 30

Glu Pro Gly Arg Arg Gly Ser Phe Val Glu Met Val Asp Asn Leu Arg
35 40 45

Gly Lys Ser Gly Gln Gly Tyr Tyr Val Glu Met Thr Val Gly Ser Pro
50 55 60

Pro Gln Thr Leu Asn Ile Leu Val Asp Thr Gly Ser Ser Asn Phe Ala
65 70 75 80

Val Gly Ala Ala Pro His Pro Phe Leu His Arg Tyr Tyr Gln Arg Gln
85 90 95

Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys Gly Val Tyr Val Pro Tyr
100 105 110

Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly Thr Asp Leu Val Ser Ile
115 120 125

Pro His Gly Pro Asn Val Thr Val Arg Ala Asn Ile Ala Ala Ile Thr

130 135 140

Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser Asn Trp Glu Gly Ile Leu

145 150 155 160

Gly Leu Ala Tyr Ala Glu Ile Ala Arg Pro Asp Asp Ser Leu Glu Pro

165 170 175

Phe Phe Asp Ser Leu Val Lys Gln Thr His Val Pro Asn Leu Phe Ser

180 185 190

Leu His Leu Cys Gly Ala Gly Phe Pro Leu Asn Gln Ser Glu Val Leu

195 200 205

Ala Ser Val Gly Gly Ser Met Ile Ile Gly Gly Ile Asp His Ser Leu

210 215 220

Tyr Thr Gly Ser Leu Trp Tyr Thr Pro Ile Arg Arg Glu Trp Tyr Tyr

225 230 235 240

Glu Val Ile Ile Val Arg Val Glu Ile Asn Gly Gln Asp Leu Lys Met

245 250 255

Asp Cys Lys Glu Tyr Asn Tyr Asp Lys Ser Ile Val Asp Ser Gly Thr

260 265 270

Thr Asn Leu Arg Leu Pro Lys Lys Val Phe Glu Ala Ala Val Lys Ser

275 280 285

Ile Lys Ala Ala Ser Ser Thr Glu Lys Phe Pro Asp Gly Phe Trp Leu

290 295 300

Gly Glu Gln Leu Val Cys Trp Gln Ala Gly Thr Thr Pro Trp Asn Ile

305 310 315 320

Phe Pro Val Ile Ser Leu Tyr Leu Met Gly Glu Val Thr Asn Gln Ser

325 330 335

Phe Arg Ile Thr Ile Leu Pro Gln Gln Tyr Leu Arg Pro Val Glu Asp

340 345 350

Val Ala Thr Ser Gln Asp Asp Cys Tyr Lys Phe Ala Ile Ser Gln Ser

355 360 365

Ser Thr Gly Thr Val Met Gly Ala Val Ile Met Glu Gly Phe Tyr Val

370 375 380

Val Phe Asp Arg Ala Arg Lys Arg Ile Gly Phe Ala Val Ser Ala Cys

385 390 395 400

His Val His Asp Glu Phe Arg Thr Ala Ala Val Glu Gly Pro Phe Val

405 410 415

Thr Leu Asp Met Glu Asp Cys Gly Tyr Asn Ile Pro Gln Thr Asp Glu

420 425 430

Ser

<210> 27

<211> 1278

<212> DNA

<213> Homo sapiens

<400> 27

atggctagca tgactggtgg acagcaaatg ggtcgcggat cgatgactat ctctgactct 60

ccgctggact ctggtatcga aaccgacgga tcctttgtgg agatggtgga caacctgagg 120

ggcaagtcgg ggcagggcta ctacgtggag atgaccgtgg gcagccccc gcagacgctc 180

aacatcctgg tggatacagg cagcagtaac tttgcagtgg gtgctgcccc ccaccccttc 240

ctgcatcgct actaccagag gcagctgtcc agcacatacc gggacctccg gaagggtgtg 300

tatgtgccct acacccaggg caagtgggaa ggggagctgg gcaccgacct ggtaagcatc 360

ccccatggcc ccaacgtcac tgtgcgtgcc aacattgctg ccatcactga atcagacaag 420

ttcttcatca acggctccaa ctgggaaggc atcctggggc tggcctatgc tgagattgcc 480

aggcctgacg actccctgga gcctttcttt gactctctgg taaagcagac ccacgttccc 540

aacctcttct ccctgcacct ttgtggtgct ggcttccccc tcaaccagtc tgaagtgctg 600

gcctctgtcg gagggagcat gatcattgga ggtatcgacc actcgctgta cacaggcagt 660

ctctggtata cacccatccg gcgggagtgg tattatgagg tcatcattgt gcgggtggag 720

atcaatggac aggatctgaa aatggactgc aaggagtaca actatgacaa gagcattgtg 780

gacagtggca ccaccaacct tcgtttgccc aagaaagtgt ttgaagctgc agtcaaatcc 840

atcaaggcag cctcctccac ggagaagttc cctgatggtt tctggctagg agagcagctg 900

gtgtgctggc aagcaggcac cacccccttgg aacatttcc cagtcatctc actctaccta 960

atgggtgagg ttaccaacca gtccttccgc atcaccatcc ttccgcagca atacctgcgg 1020

ccagtggaag atgtggccac gtcccaagac gactgttaca agtttgccat ctcacagtca 1080

tccacgggca ctgttatggg agctgttatc atggagggct tctacgttgt ctttgatcgg 1140

gcccgaaaac gaattggctt tgctgtcagc gcttgccatg tgcacgatga gttcaggacg 1200

gcagcggtgg aaggcccttt tgtcaccttg gacatggaag actgtggcta caacattcca 1260

cagacagatg agtcatga 1278

<210> 28

<211> 425

<212> PRT

<213> Homo sapiens

<400> 28

Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Gly Ser Met Thr
1   5   10   15

Ile Ser Asp Ser Pro Leu Asp Ser Gly Ile Glu Thr Asp Gly Ser Phe
            20   25   30

Val Glu Met Val Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr
        35   40   45

Val Glu Met Thr Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val
    50   55   60

Asp Thr Gly Ser Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe
65   70   75   80

Leu His Arg Tyr Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu
                85   90   95

Arg Lys Gly Val Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu
            100   105   110

Leu Gly Thr Asp Leu Val Ser Ile Pro His Gly Pro Asn Val Thr Val
        115   120   125

Arg Ala Asn Ile Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn
    130   135   140

Gly Ser Asn Trp Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala

145 150 155 160

Arg Pro Asp Asp Ser Leu Glu Pro Phe Phe Asp Ser Leu Val Lys Gln

165 170 175

Thr His Val Pro Asn Leu Phe Ser Leu His Leu Cys Gly Ala Gly Phe

180 185 190

Pro Leu Asn Gln Ser Glu Val Leu Ala Ser Val Gly Gly Ser Met Ile

195 200 205

Ile Gly Gly Ile Asp His Ser Leu Tyr Thr Gly Ser Leu Trp Tyr Thr

210 215 220

Pro Ile Arg Arg Glu Trp Tyr Tyr Glu Val Ile Ile Val Arg Val Glu

225 230 235 240

Ile Asn Gly Gln Asp Leu Lys Met Asp Cys Lys Glu Tyr Asn Tyr Asp

245 250 255

Lys Ser Ile Val Asp Ser Gly Thr Thr Asn Leu Arg Leu Pro Lys Lys

260 265 270

Val Phe Glu Ala Ala Val Lys Ser Ile Lys Ala Ala Ser Ser Thr Glu

275 280 285

Lys Phe Pro Asp Gly Phe Trp Leu Gly Glu Gln Leu Val Cys Trp Gln

290 295 300

Ala Gly Thr Thr Pro Trp Asn Ile Phe Pro Val Ile Ser Leu Tyr Leu

305 310 315 320

Met Gly Glu Val Thr Asn Gln Ser Phe Arg Ile Thr Ile Leu Pro Gln

325 330 335

Gln Tyr Leu Arg Pro Val Glu Asp Val Ala Thr Ser Gln Asp Asp Cys

340 345 350

Tyr Lys Phe Ala Ile Ser Gln Ser Ser Thr Gly Thr Val Met Gly Ala

355 360 365

Val Ile Met Glu Gly Phe Tyr Val Val Phe Asp Arg Ala Arg Lys Arg

370 375 380

Ile Gly Phe Ala Val Ser Ala Cys His Val His Asp Glu Phe Arg Thr

385 390 395 400

Ala Ala Val Glu Gly Pro Phe Val Thr Leu Asp Met Glu Asp Cys Gly

405 410 415

Tyr Asn Ile Pro Gln Thr Asp Glu Ser

420 425

<210> 29

<211> 1362

<212> DNA

<213> Homo sapiens

<400> 29

atggcccaag ccctgccctg gctcctgctg tggatgggcg cgggagtgct gcctgcccac 60

ggcacccagc acggcatccg gctgcccctg cgcagcggcc tgggggggcgc ccccctgggg 120

ctgcggctgc cccgggagac cgacgaagag cccgaggagc ccggccggag gggcagcttt 180

gtggagatgg tggacaacct gaggggcaag tcggggcagg gctactacgt ggagatgacc 240

gtgggcagcc ccccgcagac gctcaacatc ctggtggata caggcagcag taactttgca 300

gtgggtgctg ccccccaccc cttcctgcat cgctactacc agaggcagct gtccagcaca 360

taccgggacc tccggaaggg tgtgtatgtg ccctacaccc agggcaagtg ggaaggggag 420

ctgggcaccg acctggtaag catcccccat ggccccaacg tcactgtgcg tgccaacatt 480

gctgccatca ctgaatcaga caagttcttc atcaacggct ccaactggga aggcatcctg 540

gggctggcct atgctgagat tgccaggcct gacgactccc tggagccttt ctttgactct 600

ctggtaaagc agacccacgt tcccaacctc ttctccctgc acctttgtgg tgctggcttc 660

cccctcaacc agtctgaagt gctggcctct gtcggaggga gcatgatcat tggaggtatc 720

gaccactcgc tgtacacagg cagtctctgg tatacaccca tccggcggga gtggtattat 780

gaggtcatca ttgtgcgggt ggagatcaat ggacaggatc tgaaaatgga ctgcaaggag 840

tacaactatg acaagagcat tgtggacagt ggcaccacca accttcgttt gcccaagaaa 900

gtgtttgaag ctgcagtcaa atccatcaag gcagcctcct ccacggagaa gttccctgat 960

ggtttctggc taggagagca gctggtgtgc tggcaagcag gcaccacccc ttggaacatt 1020

ttcccagtca tctcactcta cctaatgggt gaggttacca accagtcctt ccgcatcacc 1080

atccttccgc agcaatacct gcggccagtg gaagatgtgg ccacgtccca agacgactgt 1140

tacaagtttg ccatctcaca gtcatccacg ggcactgtta tgggagctgt tatcatggag 1200

ggcttctacg ttgtctttga tcgggcccga aaacgaattg gctttgctgt cagcgcttgc 1260

catgtgcacg atgagttcag gacggcagcg gtggaaggcc cttttgtcac cttggacatg 1320

gaagactgtg gctacaacat tccacagaca gatgagtcat ga 1362

<210> 30

<211> 453

<212> PRT

<213> Homo sapiens

<400> 30

Met Ala Gln Ala Leu Pro Trp Leu Leu Leu Trp Met Gly Ala Gly Val

1 5 10 15

Leu Pro Ala His Gly Thr Gln His Gly Ile Arg Leu Pro Leu Arg Ser

20 25 30

Gly Leu Gly Gly Ala Pro Leu Gly Leu Arg Leu Pro Arg Glu Thr Asp

35 40 45

Glu Glu Pro Glu Glu Pro Gly Arg Arg Gly Ser Phe Val Glu Met Val

50 55 60

Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val Glu Met Thr

65 70 75 80

Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp Thr Gly Ser

85 90 95

Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu His Arg Tyr

100 105 110

Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys Gly Val

115 120 125

Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly Thr Asp

130 135 140

Leu Val Ser Ile Pro His Gly Pro Asn Val Thr Val Arg Ala Asn Ile

145 150 155 160

Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser Asn Trp

165 170 175

Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg Pro Asp Asp

180 185 190

Ser Leu Glu Pro Phe Phe Asp Ser Leu Val Lys Gln Thr His Val Pro

195 200 205

Asn Leu Phe Ser Leu Gln Leu Cys Gly Ala Gly Phe Pro Leu Asn Gln

210 215 220

Ser Glu Val Leu Ala Ser Val Gly Gly Ser Met Ile Ile Gly Gly Ile

225 230 235 240

Asp His Ser Leu Tyr Thr Gly Ser Leu Trp Tyr Thr Pro Ile Arg Arg

245 250 255

Glu Trp Tyr Tyr Glu Val Ile Ile Val Arg Val Glu Ile Asn Gly Gln

260 265 270

Asp Leu Lys Met Asp Cys Lys Glu Tyr Asn Tyr Asp Lys Ser Ile Val

275 280 285

Asp Ser Gly Thr Thr Asn Leu Arg Leu Pro Lys Lys Val Phe Glu Ala

290 295 300

Ala Val Lys Ser Ile Lys Ala Ala Ser Ser Thr Glu Lys Phe Pro Asp

305 310 315 320

Gly Phe Trp Leu Gly Glu Gln Leu Val Cys Trp Gln Ala Gly Thr Thr

325 330 335

Pro Trp Asn Ile Phe Pro Val Ile Ser Leu Tyr Leu Met Gly Glu Val

340 345 350

Thr Asn Gln Ser Phe Arg Ile Thr Ile Leu Pro Gln Gln Tyr Leu Arg

355 360 365

Pro Val Glu Asp Val Ala Thr Ser Gln Asp Asp Cys Tyr Lys Phe Ala

370 375 380

Ile Ser Gln Ser Ser Thr Gly Thr Val Met Gly Ala Val Ile Met Glu

385 390 395 400

Gly Phe Tyr Val Val Phe Asp Arg Ala Arg Lys Arg Ile Gly Phe Ala

405 410 415

Val Ser Ala Cys His Val His Asp Glu Phe Arg Thr Ala Ala Val Glu

420 425 430

Gly Pro Phe Val Thr Leu Asp Met Glu Asp Cys Gly Tyr Asn Ile Pro

435 440 445

Gln Thr Asp Glu Ser

450

<210> 31

<211> 1380

<212> DNA

<213> Homo sapiens

<400> 31

atggcccaag ccctgccctg gctcctgctg tggatgggcg cgggagtgct gcctgcccac 60

ggcacccagc acggcatccg gctgcccctg cgcagcggcc tgggggcgc cccctgggg 120

ctgcggctgc cccgggagac cgacgaagag cccgaggagc ccggccggag gggcagcttt 180

gtggagatgg tggacaacct gaggggcaag tcggggcagg gctactacgt ggagatgacc 240

gtgggcagcc ccccgcagac gctcaacatc ctggtggata caggcagcag taactttgca 300

gtgggtgctg ccccccaccc cttcctgcat cgctactacc agaggcagct gtccagcaca 360

taccgggacc tccggaaggg tgtgtatgtg ccctacaccc agggcaagtg ggaaggggag 420

ctgggcaccg acctggtaag catcccccat ggccccaacg tcactgtgcg tgccaacatt 480

gctgccatca ctgaatcaga caagttcttc atcaacggct ccaactggga aggcatcctg 540

gggctggcct atgctgagat tgccaggcct gacgactccc tggagccttt ctttgactct 600

ctggtaaagc agacccacgt tcccaacctc ttctccctgc acctttgtgg tgctggcttc 660

cccctcaacc agtctgaagt gctggcctct gtcggaggga gcatgatcat tggaggtatc 720

gaccactcgc tgtacacagg cagtctctgg tatacaccca tccggcggga gtggtattat 780

gaggtcatca ttgtgcgggt ggagatcaat ggacaggatc tgaaaatgga ctgcaaggag 840

tacaactatg acaagagcat tgtggacagt ggcaccacca accttcgttt gcccaagaaa 900

gtgtttgaag ctgcagtcaa atccatcaag gcagcctcct ccacggagaa gttccctgat 960

ggtttctggc taggagagca gctggtgtgc tggcaagcag gcaccacccc ttggaacatt 1020

ttcccagtca tctcactcta cctaatgggt gaggttacca accagtcctt ccgcatcacc 1080

atccttccgc agcaatacct gcggccagtg gaagatgtgg ccacgtccca agacgactgt 1140

tacaagtttg ccatctcaca gtcatccacg ggcactgtta tgggagctgt tatcatggag 1200

ggcttctacg ttgtctttga tcgggcccga aaacgaattg gctttgctgt cagcgcttgc 1260

catgtgcacg atgagttcag gacggcagcg gtggaaggcc cttttgtcac cttggacatg 1320

gaagactgtg gctacaacat tccacagaca gatgagtcac agcagcagca gcagcagtga 1380

<210> 32

<211> 459

<212> PRT

<213> Homo sapiens

<400> 32

Met Ala Gln Ala Leu Pro Trp Leu Leu Leu Trp Met Gly Ala Gly Val

1 5 10 15

Leu Pro Ala His Gly Thr Gln His Gly Ile Arg Leu Pro Leu Arg Ser

20 25 30

Gly Leu Gly Gly Ala Pro Leu Gly Leu Arg Leu Pro Arg Glu Thr Asp

35 40 45

Glu Glu Pro Glu Glu Pro Gly Arg Arg Gly Ser Phe Val Glu Met Val

50 55 60

Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val Glu Met Thr

65 70 75 80

Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp Thr Gly Ser

85 90 95

Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu His Arg Tyr

100 105 110

Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys Gly Val

115 120 125

Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly Thr Asp

130 135 140

Leu Val Ser Ile Pro His Gly Pro Asn Val Thr Val Arg Ala Asn Ile

145 150 155 160

Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser Asn Trp

165 170 175

Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg Pro Asp Asp

180 185 190

Ser Leu Glu Pro Phe Phe Asp Ser Leu Val Lys Gln Thr His Val Pro

195 200 205

Asn Leu Phe Ser Leu Gln Leu Cys Gly Ala Gly Phe Pro Leu Asn Gln

210 215 220

Ser Glu Val Leu Ala Ser Val Gly Gly Ser Met Ile Ile Gly Gly Ile

225 230 235 240

Asp His Ser Leu Tyr Thr Gly Ser Leu Trp Tyr Thr Pro Ile Arg Arg

245 250 255

Glu Trp Tyr Tyr Glu Val Ile Ile Val Arg Val Glu Ile Asn Gly Gln

260 265 270

Asp Leu Lys Met Asp Cys Lys Glu Tyr Asn Tyr Asp Lys Ser Ile Val

275 280 285

Asp Ser Gly Thr Thr Asn Leu Arg Leu Pro Lys Lys Val Phe Glu Ala

290 295 300

Ala Val Lys Ser Ile Lys Ala Ala Ser Ser Thr Glu Lys Phe Pro Asp

305 310 315 320

Gly Phe Trp Leu Gly Glu Gln Leu Val Cys Trp Gln Ala Gly Thr Thr

325 330 335

Pro Trp Asn Ile Phe Pro Val Ile Ser Leu Tyr Leu Met Gly Glu Val

340 345 350

Thr Asn Gln Ser Phe Arg Ile Thr Ile Leu Pro Gln Gln Tyr Leu Arg

355 360 365

Pro Val Glu Asp Val Ala Thr Ser Gln Asp Asp Cys Tyr Lys Phe Ala

370 375 380

Ile Ser Gln Ser Ser Thr Gly Thr Val Met Gly Ala Val Ile Met Glu

385 390 395 400

Gly Phe Tyr Val Val Phe Asp Arg Ala Arg Lys Arg Ile Gly Phe Ala

405 410 415

Val Ser Ala Cys His Val His Asp Glu Phe Arg Thr Ala Ala Val Glu

420 425 430

Gly Pro Phe Val Thr Leu Asp Met Glu Asp Cys Gly Tyr Asn Ile Pro

435 440 445

Gln Thr Asp Glu Ser His His His His His His

450 455

<210> 33

<211> 25

<212> PRT

<213> Homo sapiens

<400> 33

Ser Glu Gln Gln Arg Arg Pro Arg Asp Pro Glu Val Val Asn Asp Glu

1 5 10 15

Ser Ser Leu Val Arg His Arg Trp Lys

20 25

<210> 34

<211> 19

<212> PRT

<213> Homo sapiens

<400> 34

Ser Glu Gln Leu Arg Gln Gln His Asp Asp Phe Ala Asp Asp Ile Ser

1 5 10 15

Leu Leu Lys

<210> 35

<211> 29

<212> DNA

<213> Homo sapiens

<400> 35

gtggatccac ccagcacggc atccggctg 29

<210> 36

<211> 36

<212> DNA

<213> Homo sapiens

<400> 36

gaaagctttc atgactcatc tgtctgtgga atgttg 36

<210> 37

<211> 39

<212> DNA

<213> Homo sapiens

<400> 37

gatcgatgac tatctctgac tctccgcgtg aacaggacg 39

<210> 38

<211> 39

<212> DNA

<213> Homo sapiens

<400> 38

gatccgtcct gttcacgcgg agagtcagag atagtcatc 39

<210> 39

<211> 77

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Hu-Asp2

<400> 39

cggcatccgg ctgcccctgc gtagcggtct gggtggtgct ccactgggtc tgcgtctgcc 60

ccgggagacc gacgaag 77

<210> 40

<211> 77

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Hu-Asp2

<400> 40

cttcgtcggt ctcccggggc agacgcagac ccagtggagc accacccaga ccgctacgca 60

ggggcagccg gatgccg 77

<210> 41

<211> 51

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Caspase 8

Cleavage Site

<400> 41

gatcgatgac tatctctgac tctccgctgg actctggtat cgaaaccgac g 51

<210> 42

<211> 51

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Caspase 8

Cleavage Site

<400> 42

gatccgtcgg tttcgatacc agagtccagc ggagagtcag agatagtcat c 51

<210> 43

<211> 32

<212> DNA

<213> Homo sapiens

<400> 43

aaggatcctt tgtggagatg gtggacaacc tg 32

<210> 44

<211> 36

<212> DNA

<213> Homo sapiens

<400> 44

gaaagctttc atgactcatc tgtctgtgga atgttg 36

<210> 45

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: 6-His tag

<400> 45

gatcgcatca tcaccatcac catg 24

<210> 46

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: 6-His tag

<400> 46

gatccatggt gatggtgatg atgc 24

<210> 47

<211> 354

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Introduce KK
motif

<40C> 47

bbttaanvtt nnnnngactg accactcgac caggttcbnr macmhadata ragrahntsn 60
ayrsks0sna yrtawsddcg tmsnwrmans ymbarahr0g actgaccact cgaccaggtt 120
csnayrsnay rh0dtgactg accactcgac caggttcact snayrctcsn asnanrmadt 180
csnayrtcna mcrstwrd0t dthharmaca hngactgacc actcgaccag gttcttdgda 240
n0bd0cda00 a0ca0rtntr ygtabwrddc mntsmmaryn rmatndcmnt smmarynrma 300
tnsks0ycmb abctrhvgrr ccr0rsmcrs twrddcmntm swrddcwrdd cmnt 354

<210> 48

<211> 462

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Introduce KK
motif

<400> 48

bbttaanttn nnnkncgaat taaattccag cacactggct acttcttgtt ctgcatctca 60
aagaacbnrm acmhadatar agrahntsna yrsks0snay rtawsddcgt msnwrmansy 120
mbarahr0cg aattaaattc cagcacactg gctacttctt gttctgcatc tcaaagaacs 180
nayrsnayrh 0htcgaatta aattccagca cactggctac ttcttgttct gcatctcaaa 240
gaacgaasna yrttcsnasn anrmadtcsn ayrtcnamcr stwrd0cgks kdhharmaca 300
hncgaattaa attccagcac actggctact tcttgttctg catctcaaag aacttdgdan 360
0b0cda00a0 ca0rtntryh kktabwrddc mntsmmaryn rmatndcmnt smmarynrma 420
tntdccmbbc tckkmcrstw rddcmntmsw rddcwrddcm nt 462

<210> 49

<211> 380

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Introduce KK
motif

<400> 49

bbttaanttn nnnmncgaat taaattccag cacactggct abnrmacmha dataragrah 60
ntsnayrsks 0snayrtaws ddcgtmsnwr mansymbara hr0cgaatta aattccagca 120
cactggctas nayrsnayrh 0dhcgaatta aattccagca cactggctag aasnayrttc 180
snasnanrma dtcsnayrtc namcrstwrd 0cmdhharma cahncgaatt aaattccagc 240

```
acactggcta ttdgdan0b0 cda00a0ca0 rtntrymkmt abwrddcmnt smmarynrma 300
tndcmntsmm arynrmatns ks0ycmbmmc rbanbctkmk mg0g0gccr0 rsmcrstwrd 360
dcmntmswrd dcwrddcmnt                                             380
```

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification: C12N 15/57, C07K 14/47, C07K 16/18, C07K 19/00, C12N 1/21, C12N 5/10, C12N 9/64, C12N 15/12, C12N 15/62, C12N 15/85, C12Q 1/37, G01N 33/68	A3	(11) International Publication Number:	WO 00/17369
		(43) International Publication Date:	30 March 2000 (30.03.2000)

(21) International Application Number: PCT/US99/20881

(22) International Filing Date: 23 September 1999 (23.09.1999)

(30) Priority Data:
60/101,594 24 September 1998 (24.09.1998) US

(60) Parent Application or Grant
PHARMACIA & UPJOHN COMPANY [/]; (). GURNEY, Mark, E. [/]; (). BIENKOWSKI, Michael, Jerome [/]; (). HEINRIKSON, Robert, Leroy [/]; (). PARODI, Luis, A. [/]; (). YAN, Riqiang [/]; (). GURNEY, Mark, E. [/]; (). BIENKOWSKI, Michael, Jerome [/]; (). HEINRIKSON, Robert, Leroy [/]; (). PARODI, Luis, A. [/]; (). YAN, Riqiang [/]; (). WOOTTON, Thomas, A. ; ().

Published

(54) Title: ALZHEIMER'S DISEASE SECRETASE
(54) Titre: SECRETASE DE LA MALADIE D'ALZHEIMER

(57) Abstract

The present invention provides the enzyme and enzymatic procedures for cleaving the 'beta' secretase cleavage site of the APP protein and associated nucleic acids, peptides, vectors, cells and cell isolates and assays.

(57) Abrégé

La présente invention porte sur l'enzyme et les procédures enzymatiques de clivage du site de clivage de la 'beta' secrétase de la protéine APP et des acides nucléiques, des peptides, des vecteurs, des cellules et des isolats cellulaires associés, et sur des dosages.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/57, 15/62, 15/85, 5/10, 9/64, C07K 19/00, 14/47, C12N 15/12, C07K 16/18, C12Q 1/37, G01N 33/68, C12N 1/21	A3	(11) International Publication Number: **WO 00/17369** (43) International Publication Date: 30 March 2000 (30.03.00)

(21) International Application Number: PCT/US99/20881

(22) International Filing Date: 23 September 1999 (23.09.99)

(30) Priority Data:
60/101,594 24 September 1998 (24.09.98) US

(71) Applicant (*for all designated States except US*): PHARMACIA & UPJOHN COMPANY [US/US]; 301 Henrietta Street, Kalamazoo, MI 49001 (US).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): GURNEY, Mark, E. [US/US]; 910 Rosewood Avenue, S.E., Grand Rapids, MI 49506 (US). BIENKOWSKI, Michael, Jerome [US/US]; 3431 Hollow Wood, Portage, MI 49024 (US). HEINRIKSON, Robert, Leroy [US/US]; 81 South Lake Doster Drive, Plainwell, MI 49080 (US). PARODI, Luis, A. [US/SE]; Grevgafan 24, S-115 43 Stockholm (SE). YAN, Riqiang [US/US]; 5026 Queen Victoria Street, Kalamazoo, MI 49009 (US).

(74) Agent: WOOTTON, Thomas, A.; Pharmacia & Upjohn Company, Intellectual Property Legal Services, 301 Henrietta Street, Kalamazoo, MI 49001 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(88) Date of publication of the international search report: 23 November 2000 (23.11.00)

(54) Title: ALZHEIMER'S DISEASE SECRETASE

(57) Abstract

The present invention provides the enzyme and enzymatic procedures for cleaving the β secretase cleavage site of the APP protein and associated nucleic acids, peptides, vectors, cells and cell isolates and assays.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/20881

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/57 C12N15/62 C12N15/85 C12N5/10 C12N9/64 C07K19/00 C07K14/47 C12N15/12 C07K16/18 C12Q1/37 G01N33/68 C12N1/21

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED
Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K C12Q G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, STRAND, WPI Data, BIOSIS, CHEM ABS Data, MEDLINE, EMBL

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 848 062 A (SMITHKLINE-BEECHAM CORPORATION) 17 June 1998 (1998-06-17) cited in the application page 2, line 10 -page 3, line 40 page 4, line 20 - line 33 page 5, line 8 - line 20 page 8, line 1 -page 9, line 25; tables 1,2	1-3, 5-21,24, 25, 28-31, 34, 37-47, 49-64, 66-69, 72-75, 77, 80-91, 95-97, 114-129, 140,141

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
26 July 2000	02.08.00

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Montero Lopez, B

3

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/20881

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	page 10, line 28 - line 44 page 11, line 10 -page 12, line 8 ---	
X	EP 0 855 444 A (SMITHKLINE-BEECHAM P.L.C.) 29 July 1998 (1998-07-29) cited in the application page 2, line 8 -page 3, line 44 page 5, line 3 - line 15 page 5, line 49 -page 6, line 3; tables 1,2 page 7, line 34 - line 50 page 10, line 20 -page 11, line 1 page 12, line 1 - line 19 page 12, line 45 -page 13, line 44 ---	1-3, 5-21,24, 25, 28-31, 34, 37-47, 49-64, 66-69, 72-75, 77, 80-91, 95-97, 114-129, 140,141
X	WO 96 40885 A (ATHENA NEUROSCIENCES) 19 December 1996 (1996-12-19) page 3, line 1 -page 5, line 26 page 8, line 1 - line 34 page 14, line 19 -page 17, line 22 page 23, line 31 -page 25, line 20 page 28, line 7 -page 48, line 13 --- -/--	1-4,6,7, 9,10, 12-21, 24,25, 28-31, 34, 37-47, 49,50, 52,53, 55-63, 67,68, 72-75, 77, 80-90, 108-129, 136-139, 141

3

INTERNATIONAL SEARCH REPORT

Internal Application No
PCT/US 99/20881

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 26059 A (ATHENA NEUROSCIENCES, INC.) 18 June 1998 (1998-06-18) page 2, line 35 -page 4, line 3 page 5, line 9 -page 11, line 5 page 11, line 10 -page 22, line 3	1-4,6,7, 9,10, 12-21, 24,25, 28-31, 34, 37-47, 49,50, 52,53, 55-63, 67,68, 72-75, 77, 80-90, 108-129, 136-139, 141
P,X	WO 99 34004 A (CHIRON CORPORATION) 8 July 1999 (1999-07-08) page 7, line 19 -page 8, line 9 page 11, line 22 -page 14, line 24 page 16, line 26 -page 21, line 1 page 21, line 20 -page 23, line 13; figure 2; examples 2,3	1-4,6,7, 9-20,24, 28-31, 34, 37-47, 49,50, 52-63, 67,68, 72-75, 77, 80-92, 95-98, 101-103, 106,107, 114-117, 120,141

-/--

3

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

International Application No
PCT/US 99/20881

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 99 46281 A (GENENTECH, INC.) 16 September 1999 (1999-09-16) page 15, line 10 - line 23 page 65, line 5 - line 25 page 130, line 30 - line 35 page 149, line 3 -page 155, line 6 page 160, line 20 - line 22 page 173, line 35 -page 175, line 23; figures 72,73; examples 32,99-107	1-4,6,7, 9-12, 18-20, 24, 28-31, 34,37, 38, 40-47, 49,50, 52-54, 61-63, 67,68, 72-75, 77,80, 81, 84-92, 95-98, 101-103, 106,107, 114-118, 120-128, 140,141
A	US 5 795 963 A (MICHAEL JOHN MULLAN) 18 August 1998 (1998-08-18) column 3, line 58 -column 6, line 21	130-135, 141
T	YAN RIQIANG ET AL.: "Membrane-anchored aspartyl protease with Alzheimer's disease beta-secretase activity" NATURE, vol. 402, 2 December 1999 (1999-12-02), pages 533-537, XP002136300 LONDON GB	

3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 99/20881

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [ ] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. [X] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. [X] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[X] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: claims 32, 33, 35, 36, 70, 71, 76, 78 and 79 and partially claims 1, 18, 28, 44, 61, 72 and 141

Present claims 1, 18, 28, 44, 61, 72 and 141 relate to an extremely large number of possible products. In fact, the claims encompass so many possible compounds that a lack of clarity (and/or conciseness) within the meaning of Article 6 PCT arises to such an extent as to render a meaningful search of the claims impossible.

Moreover, in view of the large number and also the wording of the claims presently on file, which renders it difficult, if not impossible, to determine the matter for which protection is sought, the present application fails to comply with the clarity and conciseness requirements of Article 6 PCT (see also Rule 6.1(a) PCT) to such an extent that a meaningful search is impossible.

In addition, the obscure definition of claims 32, 33, 35, 36, 70, 71, 76, 78 and 79, relating to an unidentified SEQ ID. and referring to the examples renders as well the search of these claims impracticable.

Consequently, the search has been carried out for those parts of the application which do appear to be clear, namely the particular sequences SEQ ID NOs.: 1, 2, 3, 4, 5, 6, and 8, variants, and uses thereof

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-31, 34, 37-69, 72-75, 77, 80-129, 136-140 and partially 141

   Proteases capable of cleaving the beta secretase cleavage site of APP, variants thereof; polynucleotides encoding them; vectors and host cells comprising the same; antibodies for the polypeptides and uses of the foregoing in screening tests.

2. Claims: 130-135 and partially 141

   APP isoform wherein the last two carboxy terminus amino acids are Lysine residues.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 99/20881

Patent document cited in search report		Publication date	Patent family member(s)			Publication date
EP 848062	A	17-06-1998	JP	11069981	A	16-03-1999
			US	6025180	A	15-02-2000
EP 855444	A	29-07-1998	CA	2221686	A	28-07-1998
			JP	10327875	A	15-12-1998
			JP	2000060579	A	29-02-2000
WO 9640885	A	19-12-1996	US	5744346	A	28-04-1998
			AU	6383396	A	30-12-1996
			EP	0871720	A	21-10-1998
			JP	11507538	T	06-07-1999
			US	5942400	A	24-08-1999
WO 9826059	A	18-06-1998	AU	1684097	A	03-07-1998
WO 9934004	A	08-07-1999	AU	1726199	A	19-07-1999
			AU	2014899	A	19-07-1999
			WO	9933963	A	08-07-1999
WO 9946281	A	16-09-1999	AU	3072199	A	27-09-1999
			AU	3075099	A	11-10-1999
			WO	9947677	A	23-09-1999
			AU	1532499	A	15-06-1999
			WO	9927098	A	03-06-1999
			AU	3757099	A	08-11-1999
			WO	9954467	A	28-10-1999
			AU	1070399	A	10-05-1999
			WO	9920756	A	29-04-1999
US 5795963	A	18-08-1998	US	5455169	A	03-10-1995